

19 16
SEARCH REQUEST FORM

Examiner # (Mandatory): _____ Requester's Full Name: _____

Art Unit _____ Location (Bldg/Room#): 10D11 Phone (circle 305 306 308) _____

Serial Number: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

Title of Invention _____

Inventors (please provide full names): _____

Earliest Priority Date: _____

Keywords (include any known synonyms registry numbers, explanation of initialisms):
_____**Search Topic:**

Please write detailed statement of the search topic, and the concept of the invention. Describe as specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples of relevant citations, authors, etc., if known. You may include a copy of the abstract and the broadcast or most relevant claim(s).

STAFF USE ONLY

Searcher: <u>S. Johnson</u>	Type of Search	Vendors (include cost where applicable)
Searcher Phone #: <u>315-544-9701</u>	<input type="checkbox"/> N.A. Sequence	<input type="checkbox"/> STN
Searcher Location: _____	<input type="checkbox"/> A.A. Sequence	<input type="checkbox"/> Questel/Orbit
Date Picked Up: _____	<input type="checkbox"/> Structure (#)	<input type="checkbox"/> Lexis/Nexis
Date Completed: <u>9/22/95</u>	<input type="checkbox"/> Bibliographic	<input type="checkbox"/> WWW/Internet
Clerical Prep Time: _____	<input type="checkbox"/> Litigation I	<input type="checkbox"/> In-house sequence systems (list)
Terminal Time: _____	<input type="checkbox"/> Fulltext	<input type="checkbox"/> Dialog
Number of Databases: _____	<input type="checkbox"/> Procurement	<input type="checkbox"/> Dr. Link
	<input type="checkbox"/> Other	<input type="checkbox"/> Westlaw
		<input type="checkbox"/> Other (specify) _____

Tue Sep 21 09:52:53 1999

US-09-084-491A-1.rge

Page 2

Best Available Copy

file copy

* 8316 8332: gap of unknown length
* 8333 10032: contig of 1700 bp in length
* 10033 10049: gap of unknown length
* 10050 11755: contig of 1706 bp in length
* 11756 11772: gap of unknown length
* 11773 20315: contig of 8543 bp in length
* 20316 20332: gap of unknown length
* 20333 38250: contig of 17918 bp in length
* 38251 38267: gap of unknown length
* 38268 115442: contig of 77175 bp in length
* 115443 115459: gap of unknown length
* 115460 198605: contig of 83146 bp in length.
FEATURES source
1. .198605
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DJ400N23"

BASE COUNT 53938 a 46957 c 46811 g 50746 t 153 others

ORIGIN

Query Match 68.3%; Score 1590; DB 19; Length 198605;
Best Local Similarity 99.7%; Pred. No. 0.00e+00;
Matches 1605; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
Db 125463 TATAGAGAAATGCCCTTATAGAAAGTAGAACCGTAATAATTCTCTCCAGCATCAC 125522
Cp 2316 TATAGAGAAATGCCCTTATAGAAAGTAGAACCGTAATAATTCTCTCCAGCATCAC 2257
Db 125523 TAACACCAAGAACCCACTGAGGTCTAGGTCCCCAACAGATGGCTCATAGAAAGCCC 125582
Cp 2256 TAACACCAAGAACCCACTGAGGTCTAGGTCCCCAACAGATGGCTCATAGAAAGCCC 2197
Db 125643 AGGAGCTGCCAAGGCCCTCTTGAGTGTGACATTCACAGGTTAGAGGGCTCATGGCT 125642
Cp 2196 CACTAACCCCTGTCCTCCACATGGGCACTGGCAAGGGTCTGGAAAGGAGCTCATGGCT 2137
Db 125703 GCACAGTGCACTGGGAAATGTTGGCTCCAGGAAGCTGGCTGAGCTGCTGCTGCTA 125702
Cp 2076 GCACAGTGCACTGGGAAATGTTGGCTCCAGGAAGCTGGCTGAGCTGCTGCTGCTA 2077
Db 125763 AGCTTGCTTACGACCTTGGCAAGTCTTCTGGCTCTGGCTCATGGCT 125762
Cp 2016 AGCTTGCTTACGACCTTGGCAAGTCTGGCTCATGGCT 125882
Db 125883 CTTAATGAGGCCCTTCTTAATGGTCTTGTGCTTATGAGTTAGGGCTGTGACCCCTA 125882
Cp 1956 CTTAATGAGGCCCTTCTTAATGGTCTTGTGCTTATGAGTTAGGGCTGTGACCCCTA 1897
Db 125883 GCCCAGTGCCTTACAGGAAGGGTGTGACCTGGACGCCGAGGACACCCTGTAATCA 125942
Cp 1896 GCCCAGTGCCTTACAGGAAGGGTGTGACCTGGACGCCGAG-ACCACCTGTAATCA 1838
Db 125943 GGCCCACTGGGAGCAAGGCCCTGAAGTGTCTGAGCTAGTTTC 126002
Cp 1837 GGCCCACTGGGAGCAAGGCCCTGAAGTGTCTTATTACCTTGCCTCTGAGCTAGTTTC 1778
Db 126003 AGTTAGGCCCTTCTAGAAGCCTACAAGGCCCAAGTATAATTCCCATAAGGTTAGCTA 126062
Cp 1777 AGTTAGGCCCTTCTAGAAGCCTACAAGGCCCAAGTATAATTCCCATAAGGTTAGCTA 1718
Db 126063 ACTACTCTTACCTGTGCAAAATGAAGACAGGGTCACTATCCATCCCTGAAGAGAAACTG 126122
Cp 1717 ACTACTCTTACCTGTGCAAAATGAAGACAGGGTCACTATCCATCCCTGAAGAGAAACTG 1658
Db 126123 GGGCATCCAGGTCTGCCCTCTGTCAGGCCAAAGGGGAGGAGATCTCTATCC 126182
Cp 1657 GGGCATCCAGGTCTGCCCTCTGTCAGGCCAAAGGGGAGGAGATCTCTATCC 1598

RESULT 2 HSAC002073 128978 bp DNA PRI 12-MAY-1997
LOCUS Human PAC clone DJ515N1 from 22q11.2-q22, complete sequence.
DEFINITION Human PAC clone AC002073
ACCESSION AC002073
NID 92078469
VERSION AC002073.1 GI:2078469
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
Homo.
REFERENCE 1 (bases 1 to 128978)

AUTHORS Du, Z, Scheet, P and Harper, M.
 TITLE The sequence of *H. sapiens* PAC clone DJ515N1
 JOURNAL Unpublished (1997)
 REFERENCE 2 (bases 1 to 128978)
 AUTHORS Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (12-MAY-1997)
 COMMENT Genome Sequencing Center
 Department of Genetics
 Washington University
 St. Louis MO 63108, USA
<http://genome.wustl.edu/gsc>
 mailto:sapiens@watson.wustl.edu

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequenced overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

This sequence was generated from part of bacterial clone contigs of human chromosome 22, constructed by the Sanger Centre chromosome 22 mapping group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr22/>

SOURCE INFORMATION:

This clone was derived from human PAC library RPCI-3 prepared by Pieter de Jong and coworkers at Roswell Park Cancer Institute, using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one male donor. For further details, see <http://bacpac.med.buffalo.edu/>. The clone is available from Genome Systems, Inc. (<http://www.genomesystems.com>). VECTOR: PCYPAC2

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is H_DJ400N23; the clone sequenced to the right is H_DJ412A9. Actual start of this clone is at base position 1 of H_DJ515N1.

This clone contains SITS WI-12936 (NID:g1344756) and A006121 (NID:g1341182).

FEATURES

source
 1. .128978 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="GR:2078470"
 /chromosome="22"
 /clone="DJ515N1"
 /map="22q11.2-q22"
 7. .131
 /rpt_family="L1"
 repeat_region
 223. .526
 /rpt_family="ALU"
 528. .944
 /rpt_family="L1"
 989. .1266
 /rpt_family="ALU"
 1485. .1576
 /rpt_family="L1"
 complement(1599. .1626)
 /rpt_family="MER"
 1620. .1939
 /rpt_family="MER"
 complement(1959. .2201)
 /rpt_family="MER"
 complement(20735. .20996)

repeat_region
 complement(2813. .3101)
 /rpt_family="ALU"
 complement(3383. .3639)
 /rpt_family="ALU"
 5560. .5648
 /rpt_family="ALU"
 complement(6859. .7148)
 /rpt_family="ALU"
 8165. .8447
 /rpt_family="ALU"
 complement(8453. .8742)
 /rpt_family="ALU"
 9147. .9438
 /rpt_family="ALU"
 complement(11744. .12034)
 /rpt_family="ALU"
 complement(12037. .12313)
 /rpt_family="ALU"
 12643. .12925
 /rpt_family="ALU"
 13532. .13693
 /rpt_family="ALU"
 13889. .14009
 /rpt_family="ALU"
 14446. .14613
 /note="match to human EST AA122029 (NID:g1678048)"
 zm25f10.r1
 14452. .14613
 /note="match to human EST D31562 (NID:g644442)"
 14543. .23803
 /gene="WDGSC:DJ515N1.2"
 CDS
 join(14543. .14612,15591. .15707,15803. .15922,17188. .17388)
 17494. .17572,23599. .23803
 /gene="WUGSC:DJ515N1.2"
 /note="Putative gene. Genscan predictions confirmed by EST splicing: i coded for by human cDNAs AA122029 (NID:g1678048), D31562 (NID:g644442), AA158721 (NID:g1733515), R59640 (NID:g9830335) and F13082 (NID:g709111)"
 /codon_start=1
 /protein_id="AAB54054.1"
 /db_xref="PID:92078470"
 /transl_id="GI:2078470"
 RCLNWLDQSGILASAPVSGAGNHSHCRNPDEDPRGPWCIVSGEAGVPEKPCEDLRCP
 ETTSQALPAFTTEIOEASEGPDADEVQFAPANALPARSEAAQPVIGISQRVRMNS
 KEKKDLGTLGYVLGITMMVIIIAIGAGIILGYSYKRGKDLKEQHDQKVCEREMQRITL
 PLSAFTNPCTEIVDEKTVVVHTSQTPVDPQEGETTPLMGOATGPA"
 14544. .14613
 /gene="WUGSC:DJ515N1.2"
 /note="match to human EST AA158721 (NID:g1733515)"
 z078g11.r1
 15589. .15641
 /gene="WUGSC:DJ515N1.2"
 /note="match to human EST D31562 (NID:g644442)"
 15589. .15708
 /gene="WDGSC:DJ515N1.2"
 /note="match to human EST AA158721 (NID:g1733515)"
 z078g11.r1
 15589. .15706
 /gene="WDGSC:DJ515N1.2"
 /note="match to human EST AA122029 (NID:g1678048)"
 zm25f10.r1
 complement(16489. .16778)
 /rpt_family="ALU"
 complement(18503. .18758)
 /rpt_family="ALU"
 19258. .19548
 /rpt_family="ALU"
 19569. .19858
 /rpt_family="ALU"
 complement(20735. .20996)

repeat_region
 repeat_region
 repeat_region
 repeat_region
 repeat_region
 repeat_region

```

repeat_region          22684 . 22975
  /rpt_family="ALU"
  23163 . 23289
  /rpt_family="ALU"
  23942 . 24240
  /note="match to human EST 244818 (NID:9573984)"*
  complement(25449 . 25740)
  /rpt_family="ALU"
  complement(28446 . 58170)
  /gene="LIMK-2"
  complement(28446 . 94478)
  /gene="LIMK-2"
  complement(join(28446 . 28590, 31579 . 31736, 33380 . 33435,
  34357, 35686 . 35751, 38687 . 38743, 38980 . 39111,
  39903, 40755 . 40941, 44096 . 44292, 44648 . 44753,
  46810 . 46998, 47657 . 47766, 48461 . 48596, 81068 . 81167,
  94463 . 94478))
  /note="LIMK-2"
  /note="DJ515N1.1b; match to D45906 (NID:91136921); coded
  for by human cDNA D45906 (NID:91805593)"*
  /codon_start=1
  /product="Lim kinase"
  /protein_id="AAB54056.1"
  /db_xref="GI:2078472"
  /db_xref="PID:92078472"
  /translation="MSALAGEDDWRCPPGCDHIAPSQIWYRTVNETWHGSCEFRCSSEQ
  DSITNWYYEKDGKLYCPKDYGKFGEFCHGCSILMTGPFMVAGEFKYHPECFAIMSCK
  VIEDGDAYALVQHATLYCGKCHNEVVLAPMFERLSTSEQQLPYSVTLLISMMPATE
  EVEDAISOTSOTLQLLIEHDPVSQRLDQLRILEARLAPHMQNAQGHPHALSTLDTKENLE
  GTLRRRSLSRSNSISKSPGPSSPKPEPLFLSRSDIIRSESLRCSSSYSQQIFRPQDOLIHG
  EVLGKGFQAIKVTHATGKVMVKELIRCDEETQKTFLETEVKVNRSLDHPNVLKFI
  DLNSHNCLIKLDKTVVAVDFGLSLRLEIVEKRAMEKATTKKRTRKNDRKKRYTVVG
  NPYWWAPMHLNGKSYDETVDIFSFGIVLIEIGQVYADPDLPRTDGLNVKLFWEK
  FVPTDCPAPFFPLAAICCRLEPESRPAFSKLEDSEFEALSLYLGEGLPPLPAEELDH
  CDS
  complement(join(28446 . 28590, 29751 . 30098, 31579 . 31736,
  33380 . 33435, 34183 . 34357, 35686 . 35751, 38687 . 38743,
  39980 . 39111, 39817 . 39903, 40755 . 40941, 44096 . 44292,
  44648 . 44753, 46810 . 46998, 47657 . 47766, 48461 . 48596,
  58118 . 58170))
```

Note: remainder of annotations omitted.

Query Match 67.7%; Score 1576; DB 30; Length 128978;
 Best Local Similarity 99.6%; Pred. No. 0.00e+00;
 Matches 1603; Conservative 0; Mismatches 3; Indels 4; Gaps 3;

```

Db 23597 AGGGGAAGGATTGAAGAACAGCATGATCAGAAAGTATGTGAGAGGGAGATGCCGA 23656
Qy 709 AGGGGAAGGATTGAAGAACAGCATGATCAGAAAGTATGTGAGAGGGAGATGCCGA 768
Db 23657 ATCACTGTGCCCTGTCTGCCCTACCAACCCACCTGTGAGATTGTGGATGAGACT 23716
Qy 769 ATCACTGTGCCCTGTCTGCCCTACCAACCCACCTGTGAGATTGTGGATGAGACT 828
Db 23717 GTCGTGTCCACACCAGCCAGACTCCAGTTGACCCCTACAGGGCACCACCCCCTATG
Qy 829 GGCAGGGGGACTCCTGGGCTGAGCCCCCAGTGGCAGGAGCCCATGCCAGACAC 23776
Db 23777 GGCAGGGGGACTCCTGGGCTGAGCCCCCAGTGGCAGGAGCCCATGCCAGACAC 948
Qy 889 GGCAGGGGGACTCCTGGGCTGAGCCCCCAGTGGCAGGAGCCCATGCCAGACAC 23836
Db 23837 TGGTCAGGACAGCCACCCCTACAGCTAGGAGAACCTACCTTGTTCTGGTA 1008
Qy 949 TGGTCAGGACAGCCACCCCTACAGCTAGGAGAACCTACCTTGTTCTGGTA 23896
Db 23897 AACCCCTACCACTCCCCCTACAGCTAGGAGAACCTACCTTGTTCTGGTA 23956
Qy 1009 AACCCCTACCACTCCCCCTACAGCTAGGAGAACCTACCTTGTTCTGGTA 1068
Db 24917 CAAGGCCAGCTAACCTGTGATGTCACACTGAGGCCAACATTTCTCCACTGCACTGGCTAGCAGCTCA 2086
Qy 2087 CAAGGCCAGCTAACCTGTGATGTCACACTGAGGCCAACATTTCTCCACTGCACTGGCTAGCAGCTCA 25036
Db 25037 CTTCCCTTCAGAACCCCTCCACTGCCCAATGGGAGAC -GGTTAGGGGGCTCTCA 25094
```

Best Available Copy

Oy	2147	CTTCCTTCCAGAACCCCTTCCACTGCCAATGTGGAGACAGGGTTAGTGGCTTA	2206
Db	25095	TGGAGCCATCTGCTTGGGACCTAGACCTCAGGTGGCTCTGGTTAGTGATGCTGG	25154
Oy	2207	TGGAGCCATCTGCTTGGGACCTAGACCTCAGGTGGCTCTGGTTAGTGATGCTGG	2266
Db	25155	AGAAAGAAATTACTGGAGAAATGTTTCTACTTTCTATAAAGGCATTTCCTATA	25204
Oy	2267	AGAAAGAAATTACTGGTTCTACTTTCTATAAAGGCATTTCCTATA	2316
RESULT	3	AC005478 150140 bp DNA	HTG 14-AUG-1998
LOCUS		Homo sapiens clone DJ0425I02, WORKING DRAFT SEQUENCE, 5 unordered pieces.	
DEFINITION			
ACCESSION		AC005478	
NID		93419864	
VERSION		AC005478.1 GI:3419864	
KEYWORDS		HTG; HTGS_PHASE1.	
SOURCE		human.	
ORGANISM		Homo sapiens	
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1	(bases 1 to 150140)	
AUTHORS		Waterston, R.H.	
TITLE		The sequence of Homo sapiens clone	
JOURNAL		Unpublished	
REFERENCE	2	(bases 1 to 150140)	
AUTHORS		Waterston, R.H.	
TITLE		Direct Submission	
JOURNAL		Submitted (14-AUG-1998) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA	
COMMENT		* NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.	
	1	1987: contig of 1987 bp in length	
	*	1988 2005: gap of unknown length	
	*	2006 4759: contig of 2754 bp in length	
	*	4760 4777: gap of unknown length	
	*	4778 12352: contig of 755 bp in length	
	*	12353 12370: gap of unknown length	
	*	12371 46730: contig of 34360 bp in length	
	*	46731 46748: gap of unknown length	
	*	46749 150140: contig of 103392 bp in length.	
FEATURES		Location/Qualifiers	
source	1. 150140	/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone="DJ0425I02"	
BASE COUNT	38079	a 35586 c 35793 g 40610 t 72 others	
ORIGIN			
Query Match	67.7%	Score 1576; DB 19; Length 150140;	
Best Local Similarity	99.6%	Pred. No. 0.00e+00;	
Matches 1603; Conservative	0;	Mismatches 3; Indels 4; Gaps 3;	
Db	76991	TATAGAGAAATGCCCTTATAGAAAAGTAGAACCACTGAGGTCTAGGTCCCAGCATCAC	77050
Cp	2316	TATAGAGAAATGCCCTTATAGAAAAGTAGAACCACTGAGGTCTAGGTCCCAGCATCAC	2257
Db	77051	TAACACCAAGGACCCCTGAGGTCTAGGTCCCAGCATCAC	77110
Cp	2256	TAACACCAAGGACCCCTGAGGTCTAGGTCCCAGCATCAC	2197
Db	77111	CACTAACCC--GTCTCCACATGGCAGTGGAAAGGGTCTATGGCT	77168
Cp	1118	ACAAGGAGCACTGTAGGACCTACCCAGCTTACCCCTCAGCCCCACAGGGCACCTGCTT	1059

Best Available Copy

Db	78249	CTGTCACTCTTACTAGATTGCCAAAAAGGGGGGAGTGGTAGGGTTAACGAAAC	78308	PH: 9.3	
CP	1058	CTGTCACTCTTACTAGATTGCCAAAAAGGGGGGAGTGGTAGGGTTAACGAAAC	999	Derived from dbEST (genbank accession R59582). FEATURES source	Location/Qualifiers 1. .464
Db	78309	ACAAAGTGGTAGTCCTAGCTGCACCGAGTGTCTGCAT	78368	/organism="Homo sapiens" /db_xref="taxon:9606" /map="99.7 CR from top of Chr22 linkage group"	
CP	998	ACAAAGTGGTAGTCCTAGCTGCACCGAGTGTCTGCAT	939	STS primer_bind primer_bind BASE COUNT ORIGIN	43. .167 43. .66 complement(150. .167) 109 a 108 c 119 g 125 t 3 others
Db	78369	GGGCTCTGCCACTGGGGGCTCAGGCCAGGAGTCCTAGCTGAGGGGGGG	78428	Query Match Best Local Similarity Matches	Score 403; DB 34; Length 464; Pred. No. 0.00e+00; 0; Mismatches 10; Indels 6; Gaps 6;
CP	938	GGGCTCTGCCACTGGGGGCTCAGGCCAGGAGTCCTAGCTGAGGGGGGG	879	Db 1 CP 2326	TTTTTTTTATAGAGAAATGCCTTATAGAAAAGTAGAAACCACTTAATTCCTCTCT 60 TTTTTTTTATAGAGAAATGCCTTATAGAAAAGTAGAAACCACTTAATTCCTCTCT 2267
Db	78429	GGGGTGCCTCTGAGGGTCAACTGGAGTCGGCTGGGACACAGTCTTC	78488	Db 61 CP 2266	CCAGCATCATAACCCAAGAGACCACCTGAGGTCTAGGTCCCAGATGGCTCCA 120 CCAGCATCATAACCCAAGAGACCACCTGAGGTCTAGGTCCCAGATGGCTCCA 2207
CP	878	GTGCTGCCCTCTGAGGGTCAACTGGAGTCGGCTGGGACACAGTCTTC	819	Db 121 CP 2206	TAGAAGCCCCACTAACCCCTGCTCCACATTGGGAGGTCTGGAAAGGAG 180 TAGAAGCCCCACTAACCCCTGCTCCACATTGGGAGGTCTGGAAAGGAG 2147
Db	78489	TCCACAACTCTCACAGTGGGGTGGGAGAGTGTGATTCGCTGATC	78548	Db 181 CP 2146	CTCTATGGCTAGGAGCTGGAGCTGCCAAGGGCCTCTGAGTGTGACATCACAGGTAGGGCCCTGC 240 CTCTATGGCTAGGAGCTGGAGCTGCCAAGGGCCTCTGAGTGTGACATCACAGGTAGGGCCCTGC 2087
CP	818	TCCACAACTCTCACAGTGGGGTGGGAGAGTGTGATTCGCTGATC	759	Db 241 CP 2086	TGAGCTGCTAGCACAGTGCACTGGAGAAATGTTGCAGTGGCTCAG 299 TGAGCTGCTAGCACAGTGCACTGGAGAAATGTTGCAGTGGCTCAG 2027
Db	78549	TCCCTCTCACATCATGCTGTTCAAAATCCTTCCCCCT	78598	Db 300 CP 2026	CCATGCCAGTAGCTTACGACCTGGCAAGTCACTTTCTCTGGGGTCTCG 359 CCATGCCAGTAGCTTACGACCTGGCAAGTCACTTTCTCTGGGGTCTCG 1968
CP	758	TCCCTCTCACATCATGCTGTTCAAAATCCTTCCCCCT	709	Db 360 CP 1967	TGTTCCTAAAGGTTTAATGGGTTATGGAGTTAG 419 TGT-CCAAAAGGCTTAATGAG-GCCCTTGGCAAGTCACTTTCTCTGGGGTCTCG 1911
RESULT	4			RESULT	5
LOCUS	G24591	464 bp	DNA	LOCUS	G20845
DEFINITION	human STS WI-14622.			DEFINITION	239 bp DNA
ACCESSION	G24591			ACCESSION	human STS A006121, sequence tagged site.
NID	91344917			NID	G20845
VERSION	G1:1344917			NID	91341182
KEYWORDS	STS sequence; primer; sequence tagged site.			VERSION	G20845.1
SOURCE	human STSs derived from sequences in dbEST and the Unigene collection.			KEYWORDS	GI:1341182
ORGANISM	Homo sapiens			SOURCE	STS; STS sequence; primer; sequence tagged site.
	Eukaryota; mitochondrial eukaryotes; Metazoa; Chordata;			ORGANISM	Homo sapiens
	Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.				Bukaryota; mitochondrial eukaryotes; Metazoa; Chordata;
	1 (bases 1 to 464)				Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	Hudson, T.			REFERENCE	1 (bases 1 to 239)
AUTHORS	Whitehead Institute/MIT Center for Genome Research; Physically Mapped STS			AUTHORS	Adams, M.D.
TITLE	Unpublished (1995)			JOURNAL	Unpublished (1996)
COMMENT	Contact: Thomas Hudson Whitehead Institute/MIT Center for Genome Research Whitehead Institute for Biomedical Research 9 Cambridge Center, Cambridge MA 02142 USA Tel: 617 252 1900 Fax: 617 252 1902 Email: thudson@genome.wi.mit.edu			COMMENT	Tag Adams
STOCK NUMBER				STOCK NUMBER	The Institute for Genomic Research
STOCK QUANTITY				STOCK QUANTITY	9712 Medical Center Dr., Rockville, MD 20850
STOCK UNIT				STOCK UNIT	Email: mdadams@tigr.org
STOCK UNIT				STOCK UNIT	Primer A: GTCAAGATGAGGTATATAAACAA Primer B: GTGATGTCACTCAAGA
STOCK UNIT				STOCK UNIT	MgCl2: 1.5 mM Tris-HCl: 50 mM EDTA-HCl: 10 mM

Buffer:
MgCl2: 1.5 mM
Tris-HCl: 50 mM
EDTA-HCl: 10 mM

provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). CDNA synthesis was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGAGCGGCCCTCATTTTTTTTTTTT-3']. Double-stranded CDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7R3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo.¹

¹DB_Xref="GDB:5740622"

²DB_Xref="taxon:9606"

/clone_1lib="NCI_CGAP_GCB1"

/tissue_type="germinal center B cell"
/lab_host="DH10B"
BASE COUNT 145 a 130 c 148 g 120 t

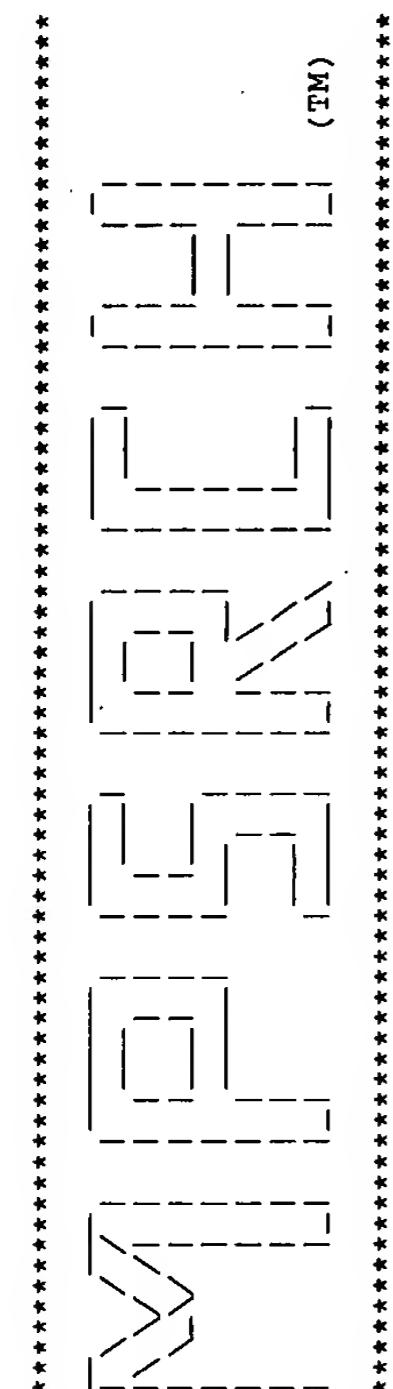
Query Match Best Local Similarity Matches	Score 22.2%; Pred. No. 0.00e+00; 0; Mismatches 0; Indels 4; Gaps 4;	Score 518; DB 10; Length 543; 0; Mismatches 0; Indels 4; Gaps 4;
Db 1 CTTCTGTTCATAAGAAGGAATTGCACATTCCACTTAGAACAGGATAG 60	QY 1544 CTTCTGTTCATAAGAAAGGAATTGCACATTCCACTTAGAACAGGATAG 1603	QY 1604 AGAAGGATCTGGCTCCGCCTTGGCACAGGAAAGGATAGAACAGGATAG 1663
Db 61 AGAAGGATCTGGCTCCGCCTTGGC-ACAGGAGCAGAGGAGACCTGGGATGCCAGCTTT 119	Db 120 CTCTCAGGGATGGATAGTGCACAGTAAGAGTAGTTAGCTA 179	Db 1664 CTCTCAGGGATGGATAGTGCCTTCATTTGCACAGTAAGAGTAGTTAGCTA 1723
QY 1624 ACCATATGGGAATTATACTGTGGGCCCTTGGCACAGGAGACCTGGGATGCCAGCTTT 163	Db 180 ACCATATGGGAATTATACTGTGGGCCCTTGGCACAGTAAGAGTAGTTAGCTA 1723	QY 1724 ACCATATGGGAATTATACTGTGGGCCCTTGTGAGCTGCTTCAACCTGGAAAC 1783
Db 240 TAAGGTCAGGGCAAGGATAAAAGCACTTCAGGGCTTGTGAGCTGGCTGTGATT 299	Db 240 TAAGGTCAGGGCAAGGATAAAAGCACTTCAGGGCTTGTGAGCTGGCTGTGATT 299	Db 240 TAAGGTCAGGGCAAGGATAAAAGCACTTCAGGGCTTGTGAGCTGGCTGTGATT 299
QY 1784 TAAGGTCAGGGCAAGGATAAAAGCACTTCAGGGCTTGTGAGCTGGCTGTGATT 1843	QY 1784 TAAGGTCAGGGCAAGGATAAAAGCACTTCAGGGCTTGTGAGCTGGCTGTGATT 1843	QY 1784 TAAGGTCAGGGCAAGGATAAAAGCACTTCAGGGCTTGTGAGCTGGCTGTGATT 1843
Db 300 AGCAGGTGGCTCTAACCTCAAAGAACATTAGAACGGCTAGGGACTGGGCTAGGGT 359	Db 300 AGCAGGTGGCTCTAACCTCAAAGAACATTAGAACGGCTAGGGACTGGGCTAGGGT 359	Db 300 AGCAGGTGGCTCTAACCTCAAAGAACATTAGAACGGCTAGGGACTGGGCTAGGGT 359
QY 1844 AGCAGGTGGCT-TGCGGGGGTCCAGGTAGGGCACTGGGACTGGGCTAGGGT 1902	QY 1844 AGCAGGTGGCT-TGCGGGGGTCCAGGTAGGGCACTGGGACTGGGCTAGGGT 1902	QY 1844 AGCAGGTGGCT-TGCGGGGGTCCAGGTAGGGCACTGGGACTGGGCTAGGGT 1902
Db 360 CACAGCCCCCTAACCTCAAAGAACATTAGAACGGCTCATTAA-CCTTT 418	Db 360 CACAGCCCCCTAACCTCAAAGAACATTAGAACGGCTCATTAA-CCTTT 418	Db 360 CACAGCCCCCTAACCTCAAAGAACATTAGAACGGCTCATTAA-CCTTT 418
QY 1903 CACAGCCCCCTAACCTCAAAGAACATTAGAACGGCTCATTAA-CCTTT 1962	QY 1903 CACAGCCCCCTAACCTCAAAGAACATTAGAACGGCTCATTAA-CCTTT 1962	QY 1903 CACAGCCCCCTAACCTCAAAGAACATTAGAACGGCTCATTAA-CCTTT 1962
Db 419 GGACACAGGCCAGAGAGAAAAAGTGACTTGGCCAAGGTGTAAGCAAGGTACTGGC 478	Db 419 GGACACAGGCCAGAGAGAAAAAGTGACTTGGCCAAGGTGTAAGCAAGGTACTGGC 478	Db 419 GGACACAGGCCAGAGAGAAAAAGTGACTTGGCCAAGGTGTAAGCAAGGTACTGGC 478
QY 1963 GGACACAGGCCAGAGAGAAAAAGTGACTTGGCCAAGGTGTAAGCAAGGTACTGGC 2022	QY 1963 GGACACAGGCCAGAGAGAAAAAGTGACTTGGCCAAGGTGTAAGCAAGGTACTGGC 2022	QY 1963 GGACACAGGCCAGAGAGAAAAAGTGACTTGGCCAAGGTGTAAGCAAGGTACTGGC 2022
Db 479 ATGGCAAGAGGCC-AGCTTCCCTGACGGAGGCCAACATTCTCCACTGCACGTGCTAGCAG 537	Db 479 ATGGCAAGAGGCC-AGCTTCCCTGACGGAGGCCAACATTCTCCACTGCACGTGCTAGCAG 537	Db 479 ATGGCAAGAGGCC-AGCTTCCCTGACGGAGGCCAACATTCTCCACTGCACGTGCTAGCAG 537
QY 2023 ATGGCAAGAGGCCAGAGAGAAAAAGTGACTTGGCCAAGGTGTAAGCAAGGTACTGGC 2082	QY 2023 ATGGCAAGAGGCCAGAGAGAAAAAGTGACTTGGCCAAGGTGTAAGCAAGGTACTGGC 2082	QY 2023 ATGGCAAGAGGCCAGAGAGAAAAAGTGACTTGGCCAAGGTGTAAGCAAGGTACTGGC 2082
Db 538 CTCAGC 543	Db 538 CTCAGC 543	Db 538 CTCAGC 543
QY 2083 CTCAGC 2088	QY 2083 CTCAGC 2088	QY 2083 CTCAGC 2088

RESULT 2
LOCUS N29083 483 bp mRNA EST 04-JAN-1996
DEFINITION Yx41f01.r1 Soares melanocyte 2NbHM Homo sapiens CDNA clone
ACCESSION N29083 IMAGE:264313 5', mRNA sequence.
NID 91147319
VERSION N29083.1 GI:1147319
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Best Available Copy
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo; Mammalia; 1 (bases 1 to 483)
REFERENCE Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tran,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tran,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT On May 9, 1995 this sequence version replaced gi:802409.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
High quality sequence stops: 452
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1501 Std Error: 0.00
Seq primer: T7
High quality sequence stop: 452.
Location/Qualifiers
1. .483
/organism="Homo sapiens"
/note="Vector: pT7R3D (Pharmacia) with a modified polylinker; Site-1: Not I; Site-2: Eco RI; 1st strand cDNA was primed with a Not I - Oligo(dR) primer [5'-TGTACCAATCTGAAGTGGAGGGCCAGTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7R3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P. Albino."
/db_xref="GDB:3873955"
/db_xref="taxon:9606"
/map="MMU16C3-C4_region"
/clone="IMAGE:264313"
/clone_lib="Soares melanocyte 2NbHM"
/sex="Male"
/tissue_type="melanocyte"
/lab_host="DH10B (ampicillin resistant)"
/base_count 108 a 105 c 139 g 130 t 1 others
ORIGIN

Query Match 20.2%; Score 471; DB 32; Length 483;
Best Local Similarity 99.4%; Pred. No. 0.00e+00;
Matches 479; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
Db 1 GGCTGAGGGTAAGGCTGGGAGCAGTGGCAGCTCCCTGGAGCAGA 60
QY 1076 GGCTGAGGGTAAGGCTGGGCTCTAACAGTGGCTCCATCCCTGGAGCAGA 1135
Db 61 TTTTGTCTGGGATGGAGACAGTGGCAGCTCCACAGTGTGCTAGGGCTTC 120
QY 1136 TTTTGTCTGGGATGGAGACAGTGGCAGCTCCACAGTGTGCTAGGGCTTC 1314
Db 121 AAACATTGGCTGCACCCCTGGAAACTGAACAGGGATAGACAGGGAGCTCCCCCAGGGCTCC 180
QY 1196 AAACATTGGCTGCACCCCTGGAAACTGAACAGGGATAGACAGGGAGCTCCCCCAGGGCTCC 1255
Db 181 TCTGTGCTTTACTAAGATGGCCTCAGTCTCCACTGTGGCTTGAAGGGCATACGTGTTA 240
QY 1256 TCTGTGCTTTACTAAGATGGCCTCAGTGTGGCTTGAAGGGCATACGTGTTA 1314
Db 241 TTCAATGGTTAAGGGTAAAGGGATGGCATTTGAAAATATTTAGTTTTAA 300
QY 1315 TTCAATGGTTAAGGGTAAAGGGATGGCATTTGAAAATATTTAGTTTTAA 1374

BASE COUNT	151	a	139	c	147	g	159	t	1 others
ORIGIN									
b	361	TCAGAACTTGGATGGAAATGAGATCTAGGTGCTGGCTGCTGGTATGCTTCAGCTTGACTGAAAG 360							
	1375	AATATTGGATGGAACTCCCTACTGACCTGGAAACTCCCTACTGACCTGGATGGAAATGAGATCTAGCTTGACTGAAAG 1434							
b	421	GGCAATGATGGCCTTGACAACCCTGGCCAGGGACTCTTCCCTGACCTGGATCTAGGTGCTGCTGGTATGCTTCAGCTTGCT 420							
	1435	TCAGAACTTGGATGGAACTCCCTACTGACCTGGATGGAAATGAGATCTAGGTGCTGCTGGTATGCTTCAGCTTGACTGAAAG 1494							
b	481	AT 482							
		Y 1555 AT 1556							
RESULT									
XCOS	N56924	597 bp mRNA EST 22-FEB-1996							
DEFINITION	YY82d03.s1	Soares_multiple_sclerosis_2NDHMSP Homo sapiens cDNA							
CLONE IMAGE	280037 3'	, mRNA sequence.							
SEQUENCE	N56924								
PARTITION	91200814								
VERSION	N56924.1	GI:1200814							
KEYWORDS	EST.								
SOURCE	human.								
ORGANISM	Homo sapiens	Eukaryota; Chordata; Craniata; Vertebrata; Mammalia;							
REFERENCE	1	(bases 1 to 597)	Eutheria; Primates; Catarrhini; Homidae; Homo.						
AUTHORS	Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.								
JOURNAL	Unpublished (1995)								
COMMENT	On Apr 14, 1993 this sequence version replaced gi:692843.								
Contact:	Wilson RK								
	Washington University School of Medicine								
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108								
TELEPHONE	314 286 1800								
FAX	314 286 1810								
EMAIL	est@watson.wustl.edu								
NOTE	This clone is available royalty-free through LInI ; contact the IMAGE Consortium (info@image.lnl.gov) for further information.								
SEQ PRIMER	m13 -40 forward								
QUALIFIERS	High quality sequence stop: 358.								
FEATURES	Location/Qualifiers								
source	1..597	/organism="Homo sapiens"							
		/note="Vector: PT7T3D (Pharmacia) with a modified polylinker V-TYPE: Phagemid; Site_1: Not I; Site_2: Eco RI; 1st strand CDNA was primed with a Not I - Oligo(dT) primer [5', TGTACCAATCTGAAGTGGGAGGCCGCATTTTTTTTT 3'], double-stranded CDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified PT7T3 vector (Pharmacia). Library went through one round of normalization to a COT = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. RNA from 4 multiple sclerosis lesions from one patient was kindly provided by Dr. Kevin G. Becker (NINDS/NIH).							
		/db_xref="GDB:3898413"							
		/db_xref="taxon:9606"							
		/clone="IMAGE:280037"							
		/clone_id="Soares_multiple_sclerosis_2NDHMSP"							
		/sex="male"							
		/tissue_type="multiple sclerosis lesions"							
		/dev_stage="Age 46"							
		/lab_host="DH10B (ampicillin resistant)"							
		CDNA LIBRARY PREPARATION: M. Bento Soares, Ph.D.							
		CDNA LIBRARY ARRAYED BY: Greg Lennon, Ph.D.							
		DNA SEQUENCING BY: Washington University Genome Sequencing Center							
		Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:							



Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (C) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
 Run on: Mon Sep 20 20:55:46 1999; MasPar time 140.28 Seconds
 1177.638 Million cell updates/sec
 Tabular output not generated.

Title: >US-09-084-491A-1
 Description: (1-2329) from US09084491A.seq
 Perfect Score: 2329
 N.A. Sequence: 1 TTACAGAACAGATAACAA.....CTCTATAAAAAAAA.....GAGATATTTTTTTTTTT
 Comp:

Scoring table: TABLE default
 Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 137044 seqs, 35465580 bases x 2

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: n-issued

1:5A_COMB 2:5B_COMB 3:5C_COMB 4:PCT9_COMB 5:backfiles1

Statistics: Mean 9.355; Variance 5.013; scale 1.866

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description	Pred. No.
C 1	64	2.7	7218	2	US-08-232-	Sequence 14, Applicati	8.57e-27
C 2	43	1.8	965	3	US-08-388-	Sequence 22, Applicati	1.40e-12
C 3	41	1.8	965	3	US-08-388-	Sequence 22, Applicati	2.70e-11
C 4	40	1.7	215	1	US-08-238-	Sequence 5, Applicatio	1.17e-10
C 5	34	1.5	7218	2	US-08-232-	Sequence 14, Applicati	6.19e-07
C 6	32	1.4	215	1	US-08-238-	Sequence 5, Applicatio	9.74e-06
C 7	28	1.2	2184	1	US-07-815-	Sequence 1, Applicatio	1.99e-03
C 8	28	1.2	2288	2	US-08-290-	Sequence 4, Applicatio	1.99e-03
C 9	28	1.2	2289	1	US-07-838-	Sequence 2, Applicatio	1.99e-03
C 10	28	1.2	2679	5	5200340-7	Patent No. 5200340.	1.99e-03
C 11	25	1.1	66	1	US-08-471-	Sequence 144, Applicati	8.77e-02
C 12	25	1.1	69	1	US-08-471-	Sequence 142, Applicati	8.77e-02
C 13	26	1.1	74	4	PCT-US95-1	Sequence 100, Applicati	2.54e-02
C 14	25	1.1	74	4	PCT-US95-1	Sequence 94, Applicati	8.77e-02
C 15	25	1.1	74	4	PCT-US95-1	Sequence 94, Applicati	8.77e-02
C 16	25	1.1	75	4	PCT-US95-1	Sequence 99, Applicati	8.77e-02
C 17	26	1.1	81	4	PCT-US95-1	Sequence 98, Applicati	2.54e-02
C 18	25	1.1	81	4	PCT-US95-1	Sequence 92, Applicati	8.77e-02
C 19	25	1.1	81	4	PCT-US95-1	Sequence 92, Applicati	8.77e-02
C 20	25	1.1	82	4	PCT-US95-1	Sequence 97, Applicati	8.77e-02

RESULT 1
 ID US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.
 AC XXXXX
 DT DE Sequence 14, Application US/08232463
 CC Sequence 14, Application US/08232463
 CC Patent No. 5670367
 GENERAL INFORMATION:
 CC APPLICANT: DORNER, F.
 CC APPLICANT: SCHEIFLINGER, F.
 CC APPLICANT: FALKNER, F. G.
 TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
 CC NUMBER OF SEQUENCES: 52
 CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Foley & Lardner
 STREET: 1800 Diagonal Road, Suite 500
 CITY: Alexandria
 STATE: VA
 COUNTRY: USA
 ZIP: 22313-0299
 COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/232,463
 CC FILING DATE:
 CC CLASSIFICATION: 435
 PRIORITY APPLICATION DATA:
 CC APPLICATION NUMBER: US/07/935,313
 CC FILING DATE:
 CC APPLICATION NUMBER: EP 91 114 300.6
 CC FILING DATE: 26-AUG-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: BENT, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 30472/114 IMMU
 TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (703) 836-9300
 CC TELEFAX: (703) 683-4109
 TELEX: 899149
 INFORMATION FOR SEQ ID NO: 14:
 CC SEQUENCE CHARACTERISTICS:

LENGTH: 7218 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 CLONE: PTZ9pt-F1S
 SEQUENCE 7218 BP; 1944 A; 1491 C; 1486 G; 1929 T; 368 OTHER.
 Query Match 2.78; Score 64; DB 2; Length 7218;
 Best Local Similarity 1.08; Pred. No. 8.57e-27;
 Matches 3; Conservative 184; Mismatches 123; Indels 0; Gaps 0;
 Sq 1064 GATYY 1123
 Db 1124 YYY 1183
 Cp 771 CCTCTGTGCACTCCCTCATACATTCTGATCATGGTCTTCATAATCCTCCC 712
 Db 1184 YY 1243
 Cp 651 GGTAATGCCAGCACGTAAGCTAGCCAGAGTTCCCAAGGTCTTGGAGTCATCCG 592
 Db 1244 YY 1303
 Cp 591 CACCCGCTGGCTGATCCAAATCACTGGCTGCCCCACTCGAGGGCAG 532
 Db 1304 YY 1363
 Cp 531 GCGGTGGCAGGAGCGAACACCTGCACCTCATCTGGACCTTCAGACGGCTTCCTG 472
 Db 1364 YYYYYYYY 1373
 Cp 471 GATTCTGTGTC 462
 RESULT²
 ID US-08-388-672A-22 STANDARD; DNA; UNC; 965 BP.
 AC xxxxxx
 DT DE Sequence 22, Application US/08388672A
 CC Sequence 22, Application US/08388672A
 CC Patent No. 5795961
 CC GENERAL INFORMATION:
 CC APPLICANT: Wallace, T. Paul
 CC APPLICANT: Harris, William J.
 CC APPLICANT: Carr, Frank J.
 CC APPLICANT: Old, Lloyd J.
 CC APPLICANT: Welt, Sydney
 CC APPLICANT: Kitamura, Kunio
 CC TITLE OF INVENTION: Recombinant Human Anti-Lewis B
 CC NUMBER OF SEQUENCES: 25
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Felfe and Lynch
 CC STREET: 805 Third Avenue
 CC CITY: New York
 CC STATE: New York
 CC COUNTRY: U.S.A.
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patentin Release #1.0, Version #1.30
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/388,672A
 CC FILING DATE: 14-FEB-1995
 CC CLASSIFICATION:
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Hanson, No. 5795961man D.
 CC REGISTRATION NUMBER: 30,946
 CC REFERENCE/DOCKET NUMBER: LUD 5409
 CC TELECOMMUNICATION INFORMATION:
 CC NAME: Hanson, No. 5795961man D.
 CC REGISTRATION NUMBER: 30,946
 CC TELEPHONE: 212-688-9200

CC TELEFAX: 212-838-3884
 CC INFORMATION FOR SEQ ID NO: 22:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 965 base pairs
 CC TYPE: nucleic acid
 CC STRANDEDNESS: unknown
 CC TOPOLOGY: unknown
 CC MOLECULE TYPE: DNA (genomic)
 SQ SEQUENCE 965 BP; 192 A; 170 C; 226 G; 200 T; 177 OTHER.

Query Match 1.8%; Score 41; DB 3; Length 965;
 Best Local Similarity 14.4%; Pred. No. 2.70e-11;
 Matches 27; Conservative 84; Mismatches 76; Indels 0; Gaps 0;

Db 776 GTTGGURHUVHSGGVRSTSTCTASDTTSYWGWRGRGWDYGGGGTYNYNGKRGRTM 835
 CP 847 GGCTGGTGTGGACCACAGCTCTCATCACAAATCTCACAGTGGGGTTGGTAAGG 788

Db 836 ADTSSNSRSVTAADTAVYYCVRGRSYDSDGDDYWGTTTVYSSHUVDKMTSSSSASYGD 895
 CP 787 CAGACAAGGGCAGAGTGATTGCTGCATCTCCCTCTCACATACTTGTATGCTGTT 728

Db 896 RVTTCRSSTTHGNNTYYWYKGKAKYRVSNRSGVSRSRGSGSGTDYTTSSDATYCGTHAR 955
 CP 727 CTTTCAAATCCTTCCCCCTCTGTAGGAGTAGGCCAGGATGATGCCAGCTCCGATGGCAA 668

Db 956 TGTKVK 962
 CP 667 TGATGAT 661

RESULT 4
 ID US-08-238-163-5 STANDARD; DNA; UNC; 215 BP.
 AC xxxxxx

DT Sequence 5, Application US/08238163
 DE Sequence 5, Application US/08238163
 CC Patent No. 5569830
 CC GENERAL INFORMATION:
 CC APPLICANT: BENNETT, Alan
 CC ADDRESS: LABAVITCH, John M.
 CC APPLICANT: POWELL, Ann
 CC APPLICANT: STOTZ, Henrik
 CC TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
 CC TITLE OF INVENTION: POLYGLACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE
 CC NUMBER OF SEQUENCES: 24
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Townsend and Townsend Khouri and Crew
 CC STREET: Steuart Street Tower, One Market Plaza
 CC CITY: San Francisco
 CC STATE: California
 CC COUNTRY: US
 CC ZIP: 94105-1493
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/238,163
 CC FILING DATE: 03-MAY-1994
 CC CLASSIFICATION: 800
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Bastian, Kevin L.
 CC REGISTRATION NUMBER: 34,774
 CC REFERENCE/DOCKET NUMBER: 2307E-540
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (415) 543-9600
 CC TELEFAX: (415) 543-5043
 CC INFORMATION FOR SEQ ID NO: 5:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 215 base pairs
 CC TYPE: nucleic acid

CC STRANDEDNESS: single
 CC TOPOLOGY: unknown
 CC MOLECULE TYPE: protein
 CC FEATURE:
 CC NAME/KEY: misc_feature
 CC LOCATION: 1..215
 CC OTHER INFORMATION: /standard_name= "Deduced amino acid
 CC OTHER INFORMATION: sequence of PGIP from bean."
 SQ SEQUENCE 215 BP; 15 A; 8 C; 25 G; 26 T; 141 OTHER.

Query Match 1.7%; Score 40; DB 1; Length 215;
 Best Local Similarity 14.9%; Pred. No. 1.17e-10;
 Matches 26; Conservative 72; Mismatches 76; Indels 1; Gaps 1;

Db 9 SSVSRTASCDNDKAKRDGNTTSSWT-TDCCNRTWGVCDTDITYRVNNDSGHNKYSSANYN 67
 CP 1005 CCAGAACACAAAGTGGTAGCTGTAGCTGTAGGAGGGTGGGCTGTGCACCAAGTG 946

Db 68 YGGNNVGAAKTHYTHHTNVSGADSKTVTDYNASGTSSSNNGTDGNRSGADSYGSSKTAM 127
 CP 945 TCTGCATGGCTCCTGCCACTGGGGCTCAGGCCAGGAGTCCGGCCTGGCCCAT 86

Db 128 TSRNRTGKTANNNAVDSRNMGDASVGSDKNTKKHAKNSADGKVGSKNNDRNNRNG 182
 CP 885 AAGGGGGTGTGCCCTCTGAGGACTAACTGGGTCAACTGGAGTCTGGCTGGACACG 831

RESULT 5
 ID US-08-232-463-14 STANDARD; DNA; UNC; 7218 BP.
 AC xxxxxx

DT Sequence 14, Application US/08232463
 DE Sequence 14, Application US/08232463
 CC Patent No. 5670367
 CC GENERAL INFORMATION:
 CC APPLICANT: SCHEIFLINGER, F.
 CC APPLICANT: FALKNER, F. G.
 CC TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
 CC NUMBER OF SEQUENCES: 52
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Foley & Lardner
 CC STREET: 1800 Diagonal Road, Suite 500
 CC CITY: Alexandria
 CC STATE: VA
 CC COUNTRY: USA
 CC ZIP: 22313-0299
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/232,463
 CC FILING DATE:
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US/07/935,313
 CC FILING DATE: 26-AUG-1991
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: BENT, Stephen A.
 CC REGISTRATION NUMBER: 29,768
 CC REFERENCE/DOCKET NUMBER: 30472/114 IMMU
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (703) 836-9300
 CC TELEFAX: (703) 683-4109
 CC TELEX: 89149
 CC INFORMATION FOR SEQ ID NO: 14:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 7218 base pairs
 CC TYPE: nucleic acid

CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC IMMEDIATE SOURCE:
 CC CLONE: PTZ9pt-F1s
 SQ SEQUENCE 7218 BP; 1944 A; 1491 C; 1486 G; 1929 T; 368 OTHER.
 Query Match 1.5%; Score 34; DB 2; Length 7218;
 Best Local Similarity 1.5%; Pred. No. 6.19e-07;
 Matches 4; Conservative 148; Mismatches 118; Indels 0; Gaps 0;
 Db 1062 GCGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1121
 |||| :
 GCGAATCACTCTGCCCTTGCTCTGAGATTGGATGAGAA 824
 QY 765 GACTGTGGTCCACACCCACCTGTGAGATTGGATGAGAA 824
 Db 1122 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1181
 ::
 QY 825 GACTGTGGTCCACACCCACCTGTGAGATTGGATGAGAA 824
 Db 1182 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1241
 ::
 QY 885 TATGGCCAGGGGACTCCTGGGCCTGAGCCCCCATGCAG 944
 Db 1242 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1301
 ::
 QY 945 ACAGTGGCAGACGCCACCCCTCTACAGCTAGGAACTACCATTGTGTTCTG 1004
 Db 1302 YYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1331
 ::
 QY 1005 GTTAACCCCTACCACTCCCCCGCTTTT 1034
 RESULT 6 ID US-08-238-163-5 STANDARD; DNA; UNC; 215 BP.
 AC xxxxxx DT Sequence 5, Application US/08238163
 CC Sequence 5, Application US/08238163
 CC Patent No. 5569830
 CC GENERAL INFORMATION:
 CC APPLICANT: BENNETT, Alan
 CC APPLICANT: LABAVITCH, John M.
 CC APPLICANT: POWELL, Ann
 CC APPLICANT: STOTZ, Henrik
 CC TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
 CC TITLE OF INVENTION: POLYGALACTURONASES AND THEIR USE TO CONTROL FUNGAL DISEASE
 CC NUMBER OF SEQUENCES: 24
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Townsend and Townsen Khourie and Crew
 CC STREET: Steuart Street Tower, One Market Plaza
 CC CITY: San Francisco
 CC STATE: California
 CC COUNTRY: US
 CC ZIP: 94105-1493
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patentin Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/238,163
 CC FILING DATE: 03-MAY-1994
 CC CLASSIFICATION: 800
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Bastian, Kevin L.
 CC REGISTRATION NUMBER: 34,774
 CC REFERENCE/DOCKET NUMBER: 2307E-540
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (415) 543-9600
 CC TELEFAX: (415) 543-5043
 CC INFORMATION FOR SEQ ID NO: 5:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 215 base pairs
 CC TYPE: nucleic acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: unknown
 CC MOLECULE TYPE: protein
 CC FEATURE: NAME/KEY: misc_feature
 CC LOCATION: 1.:215
 CC OTHER INFORMATION: /standard_name= "Deduced amino acid
 CC OTHER INFORMATION: sequence_of PGIP from bean."
 CC SEQUENCE 215 BP; 15 A; 8 C; 25 G; 26 T; 141 OTHER.
 Query Match 1.4%; Score 32; DB 1; Length 215;
 Best Local Similarity 15.5%; Pred. No. 9.74e-06;
 Matches 22; Conservative 57; Mismatches 62; Indels 1; Gaps 1;
 Db 41 WGVCDTDTTYRVNNDSGHNKYSSANYNYGGNNVGAAKTHYYTHN-VSGADS KTVTDSYN 99
 ::
 QY 629 TGGGCTACGGCTGGCATTACCATGATGGTGTATCATGCCATGGAGCTGGCATCA 688
 Db 100 ASGTSSSSNGGTDGNRSGADS YSSKTAMTSRNRTGKTTANNAVDSRNMGDASVGSDKNTKK 159
 ::
 QY 689 TCTTGGCTACTCCTACAAGAGGGAAAGGATTGAAAGAACAGCATGATCAGAAAGTAT 748
 Db 160 HAKNSADKGKVGSKNNGDRNNRY 181
 ::
 QY 749 GTGAGAGGGAGATGGCAGCGAAT 770
 RESULT 7 ID US-07-815-333A-1 STANDARD; DNA; UNC; 2184 BP.
 AC xxxxxx DT Sequence 1, Application US/07815333A
 CC Sequence 1, Application US/07815333A
 CC Patent No. 5342831
 CC GENERAL INFORMATION:
 CC APPLICANT: Nakamura, Toshikazu
 CC APPLICANT: Matsumoto, Kunio
 CC TITLE OF INVENTION: EPITHELIOCYTE GROWTH ACCELERATOR
 CC NUMBER OF SEQUENCES: 2
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Leydig, Voit & Mayer
 CC STREET: Two Prudential Plaza, Suite 4900
 CC CITY: Chicago
 CC STATE: Illinois
 CC COUNTRY: United States of America
 CC ZIP: 60601
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patentin Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/07/815,333A
 CC FILING DATE: 19911227
 CC CLASSIFICATION: 530
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Green, Robert F.
 CC REGISTRATION NUMBER: 27555
 CC REFERENCE/DOCKET NUMBER: 44069
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (312) 616-5600
 CC TELEFAX: (312) 616-5700
 CC TELEX: 25-3533
 CC INFORMATION FOR SEQ ID NO: 1:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 2184 base pairs
 CC TYPE: NUCLEIC ACID
 CC STRANDEDNESS: double
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: cDNA
 CC FEATURE: NAME/KEY: CDS
 CC LOCATION: 1.:2184

SQ SEQUENCE 2184 BP; 673 A; 455 C; 500 G; 556 T; 0 OTHER.
 Query Match Best Local Similarity 75.9%; Pred. No. 1.99e-03; Length 2184;
 Matches 41; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Db 1306 ATGAGAAATTACTGCCGAAATCCAGATGATGCTCATGGACCCCTGGTGTAC 1359
 QY 319 ATCACAGTTACTGCCGAAACCCGGACGGAGCCGGCTGGTGTAC 372

RESULT 8
 ID US-08-290-937B-4 STANDARD; DNA; UNC; 2288 BP.
 AC xxxxxx

DT Sequence 4, Application US/08290937B
 DE Sequence 4, Application US/08290937B
 CC Sequence 4, Application US/08290937B
 CC Patent No. 5648233

GENERAL INFORMATION:
 CC APPLICANT: YAMAGUCHI, KYOJI
 CC APPLICANT: SHIMA, NOBUYUKI
 CC APPLICANT: MURAKAMI, AKIHIKO
 CC APPLICANT: GOTO, MASAAKI
 CC APPLICANT: TSUDA, EISUKE
 CC APPLICANT: MASUNAGA, HIROAKI
 CC APPLICANT: TAKAHIRA, REIKO
 CC APPLICANT: OOGAKI, FUMIKO
 CC APPLICANT: UEDA, MASATSUGU
 CC APPLICANT: HIGASHIO, KANJI

TITLE OF INVENTION: MODIFIED TCF
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Testa, Hurwitz & Thibeault
 STREET: 125 High St.
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/290.937B
 FILING DATE: 19-AUG-1994
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: CAMPBELL, PAULA A.
 REGISTRATION NUMBER: 32,503
 REFERENCE/DOCKET NUMBER: FJN-022
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 248-7000
 TELEFAX: (617) 248-7100
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2288 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 SEQUENCE 2288 BP; 702 A; 491 C; 524 G; 571 T; 0 OTHER.

SQ SEQUENCE 2184 BP; 673 A; 455 C; 500 G; 556 T; 0 OTHER.
 Query Match Best Local Similarity 75.9%; Pred. No. 1.99e-03; Length 2288;
 Matches 41; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Db 1367 ATGAGAAATTACTGCCGAAATCCAGATGATGCTCATGGACCCCTGGTGTAC 1420
 QY 319 ATCACAGTTACTGCCGAAACCCGGACGGAGCCGGCTGGTGTAC 372

RESULT 10
 ID 5200340-7 STANDARD; DNA; UNC; 2902 BP.
 AC xxxxxx
 DT 01-JAN-1900

Search completed: Mon Sep 20 21:00:25 1999
Job time: 279 secs.

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Mon Sep 20 20:38:15 1999; Maspar time 499.24 Seconds
999.603 Million cell updates/
Tabular output not generated

```

Title: >US-09-084-491A-1
Description: (1-2329) from US09084491A.seq
Perfect Score: 2329
N.A. Sequence: 1 TTACCAACAGCATPACAA. .... . CTCTATAAAAAAAA 2329
Comp: AATGGTCTTGCGTATTGTT. .... . GAGATATTTTTTTTTTT

```

Scoring table: TABLE default
Gap 6

```
Nmatch STD : Dbase 0; Query 0
Searched: 271905 seqs, 107135622 bases x 2
Post-processing: Minimum Match 0%
Listing first 45 summaries
```

1:part1 2:part2 3:part3 4:part4

```

8:part6 9:part3 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40 41:part41 42:part42 43:part43
44:part44 45:part45 46:part46 47:part47 48:part48
49:part49 50:part50 51:part51 52:part52 53:part53
54:part54 55:part55 56:part56 57:part57 58:part58
59:part59 60:part60

```

卷之三

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

INTRODUCTION

Result No.	Score	Query Match	Length	DB ID	Description	Pred. NO.
1	2329	100.0	2329	60	V99636 Human tissue plasminogen-activator	0.00e+00
2	270	11.6	399	60	V99640 Human t-PALP-related	3.38e-159
3	267	11.5	472	60	V99642 Human t-PALP-related	3.88e-157
4	238	10.2	250	60	V99637 Human t-PALP-related	2.82e-137
5	235	10.1	247	60	V99638 Human t-PALP-related	3.16e-135
6	191	8.2	461	60	V99639 Human t-PALP-related	2.63e-105
7	150	6.4	334	60	V99641 Human t-PALP-related	9.67e-78
8	74	3.2	291	60	V99643 Human t-PALP-related	2.54e-28

ASSIGNMENTS

```
/*tag= a
   124.:184
/*tag= b
   185.:912
/*tag= c

sig_peptide      WO9854199-A1.
mat_peptide      03-DEC-1998.
                  27-MAY-1998; U10728.
                  28-MAY-1997; US-048000.
(HUMA-) HUMAN GENOME SCI IN
Ebner R, Moore PA, Ruben SM
WPI; 99-070207/06.
P-PSDB; W87769.

New tissue plasminogen activator
diagnosis and treatment of
Claim 2; Page 54-56; 76pp;
This DNA sequence includes
plasminogen activator-like
W87769) that is a member of
homologue of tissue plasmin
```

The 2.5 kb t-PALP message has also been detected in heart, brain, lung, placenta, liver, skeletal muscle, kidney, pancreas, spleen, thymus, prostate, testis, ovary, small intestine, colon and peripheral blood leukocytes. t-PALP cDNA that encodes mature t-PALP polypeptide is deposited as ATCC 209023. Isolated nucleic acids encoding amino acids -21 to 242, -20 to 242, 1-242, 4-63 (kringle domain) and 64-242 (protease domain) of t-PALP, or encoding epitope-bearing portions of t-PALP, are also claimed as are recombinant vectors, host cells, methods for producing t-PALP polypeptides, and related nucleic acid molecules (see V999637-43). t-PALP may be used to detect and treat disorders related to the circulatory system, and to identify agonists and antagonists of t-PALP activity. The homology between t-PALP and tPA indicates that t-PALP may be involved in the regulation of normal and abnormal clotting in e.g. stroke, deep-vein thrombosis, peripheral arterial occlusion, pulmonary embolism and myocardiothrombosis.

```

Query Match      100.0%;   Score 2329;   DB 60;   Length 2329;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 2329;   Conservative 0;   Mismatches 0;   Indels 0;
Gaps 0;

```

RESULT		2	
ID	V99640 standard; DNA; 399 BP.		
AC	V99640;		
DT	29-MAR-1999 (first entry)		
DE	Human t-PALP-related DNA clone HLMFG34R.		
KW	Tissue plasminogen activator-like protease; t-PALP; human;		
DE	circulatory system-related disorder; blood clotting; stroke;		
KW	circulatory system-related disorder; blood clotting; stroke;		
KW	thrombosis; peripheral arterial occlusion; pulmonary embolism;		
KW	myocardiothrombosis; diagnosis; therapy; ss.		
KW	Homo sapiens.		
OS	WO9854199-A1.		
PN	PD 03-DEC-1998.		
PF	PF 27-MAY-1998; U10728.		
PR	PR 28-MAY-1997; US-048000.		
PA	(HUMA-) HUMAN GENOME SCI INC.		
PI	Ebner R, Moore PA, Ruben SM;		
DR	WPI; 99-070207/06.		
PT	New tissue plasminogen activator-like protease - useful in the diagnosis and treatment of circulatory system-related disorders		
PT	Claim 20; Page 61; 76pp; English.		
PS	This DNA sequence shows homology to DNA (see V99636) coding for novel human tissue plasminogen activator-like protease (t-PALP, see W87769).		
CC	It was obtained from cDNA clone HSSAX27R. A nucleic acid molecule comprising a polynucleotide having a sequence at least 95% identical to one of 7 t-PALP related sequences (see V99637-43) is claimed. Full-length t-PALP DNA (also claimed) was discovered in a cDNA library derived from activated monocytes. The 2.5 kb t-PALP message has also been detected in heart, brain, lung, placenta, liver, skeletal muscle, kidney, pancreas, spleen, thymus, prostate, testis, ovary, small intestine, colon and peripheral blood leukocytes. Vectors, host cells and methods for producing t-PALP polypeptides are claimed. The homology between t-PALP and tissue plasminogen activator indicates that t-PALP may be involved in the regulation of normal and abnormal clotting in e.g. stroke, deep-vein thrombosis, peripheral arterial occlusion, pulmonary embolism and myocardiothrombosis.		
SQ	Sequence 472 BP; 100 A; 112 C; 149 G; 111 T;		
CC	Score 267; DB 60; Length 472;		
CC	Best Local Similarity 84.4%; Pred. No. 3.88e-157;		
CC	Matches 389; Conservative 0; Mismatches 62;		
CC	Query Match 11.5%; Best Local Similarity 84.4%; Pred. No. 3.88e-159;		
CC	Sequence 399 BP; 102 A; 74 C; 116 G; 107 T; 111 T; 11.6%; Score 270; DB 60; Length 399;		
CC	Best Local Similarity 91.0%; Pred. No. 3.38e-159;		

QY	Db	121	actaaccc--gtctccacattggcaggatgtgacatcacaggtagggccctgtgagctgtctatggctta	178
	Db	179	ggagctgccaaggccctgtgagtttgcataccatcacaggtagggccctgtgagctgtctatggctta	238
	Cp	2135	GGAGCTGCCAAGGCCTCTTGAGTTGACATCACAGGTAGGGAAAGGTCTATGGCTA	2076
	Db	239	cacagtgc 247	
	Cp	2075	CACAGTGCA 2067	
RESULT	6	V99639 standard; DNA; 461 BP.		
	AC	V99639;		
	DT	29-MAR-1999 (first entry)		
	DE	Human t-PALP-related DNA clone HAPBL24R.		
	KW	tissue plasminogen activator-like protease; t-PALP; human;		
	KW	circulatory system-related disorder; blood clotting; stroke;		
	KW	thrombosis; peripheral arterial occlusion; pulmonary embolism;		
	KW	myocardiothrombosis; diagnosis; therapy; ss.		
	OS	Homo sapiens.		
	PN	WO9854199-A1.		
	PD	03-DEC-1998.		
	PF	27-MAY-1998; U10728.		
	PR	28-MAY-1997; US-048000.		
	PA	(HUMA-) HUMAN GENOME SCI INC.		
	PI	Ebner R, Moore PA, Ruben SM;		
	DR	WPI; 99-070207/06.		
	PT	New tissue plasminogen activator-like protease - useful in the diagnosis and treatment of circulatory system-related disorders		
	PS	Claim 20; Page 61; 76pp; English.		
	CC	This DNA sequence shows homology to DNA (see V99636) coding for novel human tissue plasminogen activator-like protease (t-PALP, see W87769). It was obtained from cDNA clone HAPGT42R. A nucleic acid molecule comprising a polynucleotide having a sequence at least 95% identical to one of 7 t-PALP related sequences (see V99637-43) is claimed. Full-length t-PALP DNA (also claimed) was discovered in a cDNA library derived from activated monocytes. The 2.5 kb t-PALP message has also been detected in heart, brain, lung, placenta, liver, skeletal muscle, kidney, pancreas, spleen, thymus, prostate, testis, ovary, small intestine, colon and peripheral blood leukocytes. Vectors, host cells and methods for producing t-PALP polypeptides are claimed. The homology between t-PALP and tissue plasminogen activator indicates that t-PALP may be involved in the regulation of normal and abnormal clotting in e.g. stroke, deep-vein thrombosis, peripheral arterial occlusion, pulmonary embolism and myocardiothrombosis.		
	SQ	Sequence 334 BP; 74 A; 106 G; 72 T;		
Query Match	6.4%	Score 150; DB 60;		
	Best Local Similarity	88.0%; Pred. No. 9-67e-78;		
	Matches	249; Conservative 0; Mismatches 21; Indels 13; Gaps 10;		
Db	13	ggacaggctctgactgca-qctgggactggaggcagaqccgtc---aaggggccctcggt	68	
QY	21	GGCAGGGTCTGACTGCAAAGCTGGGACTGGAGGGACATGGCTGGAAAGCCCCAAAGGGCCAGAGCTGGTGGAAAGACGCAAGACGAACTGGCTCGGT	80	
Db	69	taaacactggcgtaatcacctgca--cga-gaggcaaggatgtgtggccctgggt	125	
QY	81	TAAACACTGGTCGTCAATCACCTCGTCAGCAACATGCTCCTAGGAA-GCCATATGGATCTGG-AGGCTGT	140	
Db	126	acaaggattccct-gtcagcaacaatgctccatggatctggaggctgt	184	
QY	141	ACAAGCATTCCTCGTCAGCAACATGCTCCTAGGAA-GCCATATGGATCTGG-AGGCTGT	198	
Db	185	ttctgggacaacggccacctgttgcgttgcggcggc-cgggccttcgg	242	
QY	199	TCTGGGACAACGGCCACCTGTACGGGAGGACCAAGCTCCCCGGCCCTCGGC	258	
Db	243	ggccttcaattgggttgcgtggcaaaggggcttgtctggcc	285	
QY	259	TGCCT-CAACTGGTGGACGGCAGGGGGCTGGCTCGGCC	300	
RESULT	8			
ID	V99643	standard; DNA; 291 BP.		
AC	V99643;			
DT	29-MAR-1999 (first entry)			
DE	Human t-PALP-related DNA clone HSSES93R.			
KW	tissue plasminogen activator-like protease; t-PALP; human;			
KW	circulatory system-related disorder; blood clotting; stroke;			
Db	190	ctggaggctgtttctggacaacggccac-tgtacccggaggaccgttccc-g-gc	246	
QY	188	CTGGAGGTGTCTGGACAAACGGCCACCTGTACGGGAGGACCAAGACCTCCCCGGCGC	247	
Db	247	cggtccctc-g-tgcctcaactggcggacggcagggtgtgg	289	
QY	128	TGTTGGGCTGGGTACAAAGCATTCCTCGTCAGCAACATGCTCCTAGGAAAGCCATATGGAT	187	

FT WO9854199-A1.
 PN 03-DEC-1998; U10728.
 PF 28-MAY-1998; US-048000.
 PR (HUMA-) HUMAN GENOME SCI INC.
 PA Ebner R, Moore PA, Ruben SM;
 PI WPI; 99-070207/06.
 DR N-PSDB; V99636.

PT New tissue plasminogen activator-like protease - useful in the diagnosis and treatment of circulatory system-related disorders

PT Claim 1; Page 56-57; 76pp; English.

CC This is the amino acid sequence of tissue plasminogen activator-like protease (t-PALP), a novel member of the serine plasminogen activator family that shares sequence homology to human tissue plasminogen activator (see W87770). The t-PALP sequence was deduced from a cDNA clone (see 999636) derived from activated monocytes. The 2.5 kb t-PALP message has also been detected in heart, brain, lung, placenta, liver, skeletal muscle, kidney, spleen, thymus, prostate, testis, ovary, small intestine, colon and peripheral blood leukocytes. Isolated nucleic acids encoding amino acids 21 to 242, -20 to 242, 1-242, 4-63 (kringle domain) and 64-242 (protease domain) of t-PALP, or encoding epitope-bearing portions of t-PALP, are also claimed, as are recombinant vectors, host cells, and methods for producing t-PALP polypeptides. t-PALP may be used to detect and treat disorders related to the circulatory system, and to identify agonists and antagonists of t-PALP activity. The homology between t-PALP and tPA indicates that t-PALP may be involved in the regulation of normal and abnormal clotting in e.g. stroke, deep-vein thrombosis, peripheral arterial occlusion, pulmonary embolism and myocardiothrombosis.

SQ Sequence 263 AA;

Query Match Score 1883; DB 39; Length 263;

Best Local Similarity 100.0%; Pred. No. 1.04e-174; Mismatches 0; Indels 0; Gaps 0;

Db - 1 MLLAWQAFIYSNMLLAEAYGSGGCCFWDNGHLYREDQTSPAPGLRCLNWLDAGSGLASAP 60
 QY 1 VSGAGNHSYCRNPDEDPRGPWCYSSGEAGVPEKRPCEDLRCPETTSQALPAFTEIQEAS 120

Db 61 VSGAGNHSYCRNPDEDPRGPWCYSSGEAGVPEKRPCEDLRCPETTSQALPAFTEIQEAS 120
 QY 61 VSGAGNHSYCRNPDEDPRGPWCYSSGEAGVPEKRPCEDLRCPETTSQALPAFTEIQEAS 120

Db 121 eggaddevqvapanalparseaavqpvigisqrvmnskekdklgtlgvlgitmmvi 180
 QY 121 EGGGADEVQVFAPANALPARSEAAVQPVIGISQRVMNSKERKDGTGTVLGITMMVI 180

Db 181 iiaigagiiqlgyskrgkdlekqhdqvcereqritplsaftnptceivdektvvht 240
 QY 181 IIAIGAGIIQLGYSKRGKDLKEQHDQVKCEREMQRITLPLSAFTNPTCEIVDEKTVVHT 240

Db 241 sotqpvdpqegstplmgqagtppga 263
 QY 241 SOTQPVDPQEGSTPLMGQAGTPGA 263

RESULT 2
 ID W72641 standard; peptide; 39 AA.
 AC W72641;
 DT 05-JAN-1999 (first entry)
 DE Nervous glia cell growth factor N-terminal peptide #2.
 KW Nervous glia cell growth factor; human; urine; secretion Promoter;
 KW choline acetyltransferase activity enhancer; nervous disease.
 OS Homo sapiens.
 FH Key
 FT Misc_difference 25
 FT /note= "unspecified"
 FT Misc_difference 29
 FT /note= "unspecified"

FT Misc_difference 29
 FT /note= "unspecified"
 FT Misc_difference 29
 FT /note= "unspecified"

PN J10265498-A.
 PD 06-OCT-1998.

PF 24-MAR-1997; 090305.
 PR 24-MAR-1997; JP-090305.
 PA (NICH-) JAPAN CHEM RES CO LTD.
 DR WPI; 98-589719/50.
 PT Nervous glia cell growth factor derived from human urine - used for treatment of nervous diseases

PS Claim 3; Fig 7; 14pp; Japanese.

CC The present invention describes nervous glia cell growth factor, which is purified from human urine by ultrafiltration, salting-out by ammonium sulphate, gel filtration, ion exchange chromatography and reversed phase polyacrylamide gel electrophoresis. Also described are: (1) a secretion promoter for the nerve growth factor of glia cell consisting of the above growth factor, an enhancer for choline acetyltransferase activity of neuron consisting of the above growth factor; and (2) DNA encoding nervous glia growth factor containing a DNA sequence coding the amino acid sequence shown by the two 39 amino acid sequences as given in W72640 and W72641, which are identical, except one starts with Tyr and the other with Ser (i.e. they are from different DNA transcripts). The glia cell growth factor can be prepared in a large amount and the factor can be used for the treatment of nervous diseases.

SQ Sequence 39 AA;

Query Match Score 256; DB 36; Length 39;

Best Local Similarity 92.3%; Pred. No. 3.64e-13; Mismatches 3; Indels 0; Gaps 0;

Db 1 sgggfwndnghlyredqtspapglrxlnwxdqsglasap 39
 QY 22 SGGCFWNDNGHLYREDQTSPAPGLRCLNWLDAGSGLASAP 60

RESULT 3

ID W72640 standard; peptide; 39 AA.
 AC W72640;
 DT 05-JAN-1999 (first entry)
 DE Nervous glia cell growth factor N-terminal peptide #1.
 KW Nervous glia cell growth factor; human; urine; secretion promoter;
 KW choline acetyltransferase activity enhancer; nervous disease.
 OS Homo sapiens.
 FH Key
 FT Misc_difference 25
 FT /note= "unspecified"
 FT Misc_difference 29
 FT /note= "unspecified"

J10265498-A.

PD 06-OCT-1998.

PF 24-MAR-1997; 090305.

PR 24-MAR-1997; JP-090305.

PA (NICH-) JAPAN CHEM RES CO LTD.
 DR WPI; 98-589719/50.
 PT Nervous glia cell growth factor derived from human urine - used for treatment of nervous diseases

PS Claim 2; Fig 6; 14pp; Japanese.

CC The present invention describes nervous glia cell growth factor, which is purified from human urine by ultrafiltration, salting-out by ammonium sulphate, gel filtration, ion exchange chromatography and reversed phase polyacrylamide gel electrophoresis. Also described are: (1) a secretion promoter for the nerve growth factor containing a DNA sequence coding the amino acid sequence shown by the two 39 amino acid sequences as given in W72640 and W72641, which are identical, except one starts with Tyr and the other with Ser (i.e. they are from different DNA transcripts). The glia cell growth factor can be prepared in a large amount and the factor can be used for the treatment of nervous diseases.

SQ Sequence 39 AA;

Query Match Score 255; DB 36; Length 39;

Best Local Similarity 90.2%; Pred. No. 4.50e-13; Mismatches 2; Indels 2; Gaps 2;

RESULT 4

ID R83959 standard; protein; 812 AA.
 AC R83959;
 DT 10-MAR-1996 (first entry)
 DE Complete mouse Plasminogen molecule.
 KW Angiotatin; plasminogen; endothelial inhibitor; therapeutic;
 KW gene therapy.
 OS Mus musculus.
 FH Key
 FT protein
 FT protein
 FT protein
 FT protein
 PN WO9529242-A1.
 PD 02-NOV-1995.
 PR 26-APR-1995; US-248629.
 PR 20-OCT-1994; US-326785.
 PA (CHIL-) CHILDRENS MEDICAL CENT.
 PI Cao Y, Folkman MJ, O'Reilly MS, Sim KL;
 DR WPI; 95-382990/49.
 PT Endothelial inhibitor Angiotatin - useful to treat angiogenic
 mediated disease esp. angiogenesis and cancer.
 PS Disclosure; Fig 1; 108pp; English.
 CC Angiotatin (see R83960) is a Plasminogen fragment starting at
 amino acid 98 of the complete Plasminogen molecule. Preferably,
 angiотatin has an amino acid sequence similar to that of the
 plasminogen fragment. Angiotatin is an endothelial inhibitor,
 which reversibly inhibits proliferation of endothelial cells and
 thereby inhibits angiogenesis. It is useful in the treatment of
 a human or animal with angiogenic mediated disease e.g. arthritis,
 macular degeneration, diabetic retinopathy or cancer. Cells
 comprising angiotatin-coding sequences are useful for gene therapy
 of primary tumors.
 SQ Sequence 812 AA;

RESULT 5

ID W07585 standard; protein; 812 AA.
 AC W07585;
 DT 24-JUN-1997 (first entry)
 DE Murine Plasminogen sequence.
 KW Angiotatin; plasminogen; kringle; angiogenesis; cancer; arthritis;
 KW macular degeneration; diabetic retinopathy.
 OS Mus musculus.
 PN WO9635774-A2.
 PD 14-NOV-1996.
 PF 26-APR-1995; US-05856.
 PR 26-APR-1995; US-605598.
 PR 08-MAR-1996; US-612788.

Query Match 9.8%; Score 184; DB 15; Length 812;
Best Local Similarity 27.8%; Pred. No. 1.03e-06;
Matches 35; Conservative 38; Mismatches 43; Indels 10; Gaps 9;

Db 377 cyqsdgqsyrgtssttigkkcqswwaamfphrhsktpenfpdaglemnycrnnpdgd-kqp 435
 QY 25 CFWDNGHLYREDQTSPPAGLRLCLNWLD--AQG-GLASAPVSGAG-NHSYCRNPDEDPRGP 80

Db 436 wcytt-dpsvr-weycnlkrcsettg-svvelpvtsqeps-gpsdsetdcmyngnqkdryg 491
 QY 81 WCYVSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASEGPGADEVQ-VFAPANALPA 139

Db 492 ktavta 497
 QY 140 RSEAAA 145

Query Match 9.8%; Score 184; DB 21; Length 812;
Best Local Similarity 27.8%; Pred. No. 1.03e-06;
Matches 35; Conservative 38; Mismatches 43; Indels 10; Gaps 9;

Db 377 cyqsdgqsyrgtssttigkkcqswwaamfphrhsktpenfpdaglemnycrnnpdgd-kqp 435
 QY 25 CFWDNGHLYREDQTSPPAGLRLCLNWLD--AQG-GLASAPVSGAG-NHSYCRNPDEDPRGP 80

Db 436 wcytt-dpsvr-weycnlkrcsettg-svvelpvtsqeps-gpsdsetdcmyngnqkdryg 491
 QY 81 WCYVSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASEGPGADEVQ-VFAPANALPA 139

Db 492 ktavta 497
 QY 140 RSEAAA 145

RESULT 6

ID R89197 standard; protein; 655 AA.
AC R89197;
DT 08-AUG-1996 (first entry)
DE Human hepatocellular growth factor single chain precursor protein.
KW Mature protein; inactive; single chain; hepatocellular growth factor;
KW liver; human; precursor; dimerisation; double chain; serine protease;
KW hepatitis.
OS Homo sapiens.
FH Key
FT peptide
FT 356 .655
FT /note= "mature peptide"
PA (MITU) MITSUBISHI CHEM. CORP.
DR WPI; 96-136206/14.
PT New preventative and therapeutic cpds contg. a 34 kd serine protease
PT - useful for treatment of liver diseases e.g. hepatitis.
PS Claim 4; Page 6-8; 8pp; Japanese.

This is the amino acid sequence of the precursor protein of the inactive
 single chain form of a hepatocellular growth factor. The mature protein
 (R89196) has a mol. wt. of around 34 kd and is derived from the 96 kd
 precursor protein. The mature protein corresponds to residues 356-655 of
 the precursor protein. The inactive single chain form of the growth
 factor is activated by dimerisation of the mature protein. The active
 protein is a serine protease which can be used for the treatment of
 hepatitis. Sequence 655 AA;

Query Match 9.6%; Score 181; DB 17; Length 655;

		/note= "putative protein contains 4 kringle domains followed by a serine protease-like domain"
FT	modified_site	57..59
FT		/label= N-glycosylation_site
FT		/note= "potential"
Db	286 cflngtgrygavstsasq1sclawsndllyqelhvdsvgaalllgphaycrnpndne 345	158..160
QY	25 CFWDNGHLYREDQTSPAPGLRCLNW----L--DAQ-SGLASAPVSGAGNHSYCRNPDEDP 77	
DDB	346 r-pwcyvvkdsals 358	
QY	78 RGPWCYVSGEAGVP 91	
RESULT	7	
ID	R53962 standard; Protein; 655 AA.	
AC	R53962;	
DT	06-JAN-1995 (first entry)	
DE	Hepatocyte growth factor converting protease.	
KW	Hepatocyte growth factor; protease; cleavage; active; inactive; precursor.	
COS	Homo sapiens.	
PN	EP-596524-A.	
PD	11-MAY-1994.	
PPE	05-NOV-1993; 117988.	
PR	05-NOV-1992; JP-296133.	
PR	20-NOV-1992; JP-312234.	
PR	20-NOV-1992; JP-312242.	
PPA	(SHIM/) SHIMOMURA T.	
PA	(MITU) MITSUBISHI KASEI CORP.	
PT	Kitamura N, Miyazawa K, Morimoto Y, Shimomura T;	
PI	Yamada K; WPI; 94-152921/19.	
DR	N-PSDB; Q63951.	
PT	Hepatocyte growth factor converting protease and precursor and gene encoding them - for producing active two chain HGF from inactive single chain HGF	
PT	Claim 12; page 21-24; 30pp; English.	
PS	Hepatocyte growth factor converting protease is capable of converting inactive single chain hepatocyte growth factor (HGF) into active two chain HGF by cleavage at a specific site.	
SQ	Sequence 655 AA;	
Query Match	9.68; Score 181; DB 10; Length 655;	
Best Local Similarity	40.5%; Pred. No. 1.88e-06;	
Matches	30; Conservative 19; Mismatches 17; Indels 8; Gaps 4;	
Db	286 cflngtgrygavstsasq1sclawsndllyqelhvdsvgaalllgphaycrnpndne 345	154
QY	25 CFWDNGHLYREDQTSPAPGLRCLNWDA-QSGLASAPVSGAG-NHSYCRNPDEDP 77	82
DDB	346 r-pwcyvvkdsals 358	
QY	78 RGPWCYVSGEAGVP 91	
RESULT	9	
ID	W14271 standard; Protein; 701 AA.	
AC	W14271;	
DT	21-JUL-1997 (first entry)	
DE	Mouse growth factor L5/3 partial cDNA clone ML5-2 encoded protein.	
KW	Mouse; growth factor; foetal; liver; probe; bovine; prothrombin; locus; polymorphism; transition; exon; intron; chromosome; kringle domain; cell growth; tumour suppressor; hepatocyte growth factor; regeneration.	
KW	Key	
OS	Mus musculus.	
FH	Location/Qualifiers	
FT	peptide 1..16	
FT		/note= "partial signal peptide sequence"
FT		/note= "amino acid residue is Gln in this position in the protein encoded by the genomic sequence (T62442); this may be due to a polymorphism"
FT	protein 17..701	
FT		/note= "mature protein"
PN	US5606029-A.	
PD	25-FEB-1997.	
PF	14-MAY-1992; 882925.	
PR	14-MAY-1992; US-882925.	
RESULT	8	
ID	R66600 standard; Protein; 701 AA.	
AC	R66600;	
DT	13-FEB-1995 (first entry)	
DE	Mouse L5/3 tumour suppressor protein.	
KW	Mouse L5/3 gene; small lung cell carcinoma; tumour suppression; chromosomal 3; 3p21; D3F15S2 locus; hepatocyte growth factor; renal cell carcinoma; von Hippel-Lindau syndrome; predisposition; kringle domain.	
KW	Mus musculus.	
OS	Location/Qualifiers	
FT	peptide 1..16	
FT		/label= signal_peptide_(16-31)
FT		/note= "putative"
FT	misc_difference 4	
FT		/label= polymorphic_site
FT		/note= "Pro corresponds to CCG codon in cDNA; in the genomic DNA, codon 19 is CAG (Gln)"
FT	protein 17..701	

PR 18-JAN-1994; US-184012.
 PA (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.
 PI Degen SJ;
 DR WPI; 97-153621/14.
 DR N-PSDB; T62441.

PT Human growth factor protein L5/3 - useful for altering cell growth,
 e.g. as tumour suppressor
 Disclosure; Column 27-32; 34pp; English.

This is the amino acid sequence encoded by the insert isolated from clone MLS-2 and constitutes part of a mouse growth factor designated L5/3 (W14272). The encoding sequence was isolated from a lambda-gt10 mouse liver cDNA library using a fragment of the corresponding human cDNA sequence (T62436) as a probe. The fragment presented here is a partial sequence which lacks 44 bp of the coding region and 94 bp of the non-coding region. A 1450 bp fragment of this sequence (nucleotides 738-2188) covering 8 amino acids of the second kringle domain and all of the third and fourth kringle domains and the serine protease-like domain, was used to screen a mouse liver genomic DNA library to isolate the full length genomic sequence (T62442). The protein can be used to alter cell growth (as a growth factor or tumour suppressor) and has similar properties to the hepatocyte growth factor that is actively involved in liver regeneration.

SQ Sequence 701 AA;

Query Match 9.4%; Score 177; DB 22; Length 701;
 Best Local Similarity 41.3%; Pred. No. 4.17e-06;
 Matches 26; Conservative 9; Mismatches 26; Indels 2; Gaps 2;

Db 95 cimdnqsyrgtvarttagglpcqawsrrfpndhkytpkngleenfcrnpgdprgpwc 154
 |: ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 25 CFWDNGHLYREDQTSPAPGLRCLNWLD-QSGLASAPVSGAG-NHSYCRNPDEDPRGWC 82

Db 155 ytt 157
 | :||| :||| :||| :||| :||| :||| :||| :||| :|||
 QY 83 YVS 85

RESULT 10
 ID R666601 standard; Protein; 716 AA.
 AC R666601;
 DT 14-FEB-1995 (first entry)
 DE Mouse L5/3 tumour suppressor protein (from genomic sequence).
 KW Mouse L5/3 gene; small lung cell carcinoma; tumour suppression;
 KW chromosome 3; 3p21; D3F15S2 locus; hepatocyte growth factor;
 KW renal cell carcinoma; Von Hippel-Lindau syndrome; predisposition;
 KW kringle domain.
 OS Mus musculus.
 FH Key
 FT peptide
 FT 1..31
 FT /label= signal_peptide
 FT /note= "putative"
 FT misc_difference 19
 FT /label= polymorphic_site
 FT /note= "Gln corresponds to CAG codon in genomic DNA;
 protein in the cDNA, codon 19 is CCG (PRO)(R666600)"
 FT 32..716
 FT /note= "putative protein contains 4 kringle domains
 followed by a serine protease-like domain"
 FT modified_site 72..74
 FT /label= N-glycosylation_site
 FT /note= "potential"
 FT modified_site 173..175
 FT /label= N-glycosylation_site
 FT /note= "potential"
 FT modified_site 305..307
 FT /label= N-glycosylation_site
 FT /note= "potential"
 FT modified_site 620..622
 FT /label= N-glycosylation_site
 FT /note= "potential"
 PN US5315000-A.
 PD 24-MAY-1994.
 PR 14-MAY-1992; 882925.
 PR 14-MAY-1992; 882925.
 PR 18-JAN-1994; US-184012.
 PA (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.
 PI Degen SJ;
 DR WPI; 97-153621/14.
 DR N-PSDB; T62442.

PT Human growth factor protein L5/3 - useful for altering cell growth,
 e.g. as tumour suppressor
 Disclosure; Column 33-42; 34pp; English.

This is the amino acid sequence of the wild type mouse growth factor designated L5/3. The protein sequence differs from that encoded by the cDNA clone (T62441) at position 19; in this sequence a Gln, in the cDNA

PR 14-MAY-1992; US-882925.
 PA (CHIL-) CHILDRENS HOSPITAL MEDICAL CENT.
 PI Degen SJ;
 DR WPI; 94-166645/20.
 DR N-PSDB; Q79727.

PT DNA from D3F15S2 locus of human chromosome 3 - encoding novel growth factor, L5/3, useful as probe for detecting pre-deposition towards cancer

PS Disclosure; Columns 33-42; 31pp; English.

The sequence of mouse genomic DNA coding for the L5/3 tumour suppressor protein is composed of 18 exons separated by 17 intervening sequences. There is only one difference found between the cDNA (Q79726) and genomic DNA (Q79727) coding sequences which results in the substitution of a Gln in the gene to a Pro in the cDNA at amino acid position 19. The putative mouse protein has the same domain structure as its human homologue with four kringle domains followed by a serine protease-like domain.

SQ Sequence 716 AA;

Query Match 9.4%; Score 177; DB 10; Length 716;
 Best Local Similarity 41.3%; Pred. No. 4.17e-06;
 Matches 26; Conservative 9; Mismatches 26; Indels 2; Gaps 2;

Db 110 cimdnqsyrgtvarttagglpcqawsrrfpndhkytpkngleenfcrnpgdprgpwc 169
 |: ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 QY 25 CFWDNGHLYREDQTSPAPGLRCLNWLD-QSGLASAPVSGAG-NHSYCRNPDEDPRGWC 82

RESULT 11
 ID W14272 standard; Protein; 716 AA.
 AC W14272;
 DT 21-JUL-1997 (first entry)
 DE Mouse growth factor L5/3 complete protein.
 KW Mouse; growth factor; foetal; liver; probe; bovine; prothrombin; locus; polymorphism; transition; exon; intron; chromosome; kringle domain; cell growth; tumour suppressor; hepatocyte growth factor; regeneration.
 OS Mus musculus.
 FH Key
 FT peptide 1..31
 FT /note= "signal peptide"
 FT misc_difference 19
 FT /note= "amino acid residue is Pro at this position
 in the protein encoded by the cDNA clone ML5-2
 (T62441); this may be due to a polymorphism"
 FT protein 32..716
 FT /note= "mature protein"
 FT modified_site 72
 FT /note= "N-linked glycosylation site"
 FT modified_site 173
 FT /note= "N-linked glycosylation site"
 FT modified_site 305
 FT /note= "N-linked glycosylation site"
 FT modified_site 624
 FT /note= "N-linked glycosylation site"
 PN US5606029-A.
 PD 25-FEB-1997.
 PR 14-MAY-1992; 882925.
 PR 14-MAY-1992; 882925.
 PR 18-JAN-1994; US-184012.
 PA (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.
 PI Degen SJ;
 DR WPI; 97-153621/14.
 DR N-PSDB; T62442.

PT Human growth factor protein L5/3 - useful for altering cell growth,
 e.g. as tumour suppressor
 Disclosure; Column 33-42; 34pp; English.

This is the amino acid sequence of the wild type mouse growth factor designated L5/3. The protein sequence differs from that encoded by the cDNA clone (T62441) at position 19; in this sequence a Gln, in the cDNA

Db 386 Cyrgngknymgnlsqtrs gltcsmwdknmedlhrhifwepdasklnenycrnpddahgp 445
 |:||| :||| :||| | :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Qy 25 CFWDNGHLYREDQTS PAPGLRCLNW ---LDA-QSGLASAPVSGAGNHSYCRNPDEDPRGP 80

Db 446 wcy-tgnplip-wdyccpisrcgdtptnsqf 475
 |:||| :||| :||| | :||| :||| :||| :||| :||| :||| :||| :|||
 Qy 81 WCYVSGEAGVPEKRPCEDLRCPETTSQALPAF 112

RESULT 15
 ID R21976 standard; Protein; 728 AA.
 AC R21976;
 DT 03-JUL-1992 (first entry)

DE Human Hepatocyte growth factor.
 KW rhHGF; beta chain; rat HGF; HBC25; HAC19; ss.
 OS Homo sapiens.

PN J0403000-A.
 PD 31-JAN-1992.
 PF 24-MAY-1990; 212818.
 PR 05-JUN-1989; JP-142697.
 PR 01-JAN-1990; JP-212818.
 PR 24-MAY-1990; JP-134487.
 PA (TOYM) TOYOB KK.
 DR WPI: 92-085905/11.
 DR N-PSDB; Q22146.

PT Recombinant human hepatocyte growth factor - for treatment and
 PT diagnosis of liver diseases
 PS Claim 2; Fig 4; 21pp; Japanese.
 CC A first cDNA library (I) was prepared from human liver RNA. The
 CC library was screened with the rat HGF beta chain coding sequence
 CC RBC1 (see Q22142). Clone HBC25 was isolated and sequenced (see
 CC Q22143). The partial human HGF coding sequence was itself used
 CC as a probe to screen a second cDNA library (II). A clone which
 CC partially overlapped with HBC25 was identified and designated
 CC HAC19 (see Q22144). The complete human HGF coding sequence could
 CC then be derived from the two overlapping sequences and the amino
 CC acid sequence of HGF deduced from it.
 CC See also Q22141 and Q22145.
 SQ Sequence 728 AA;

Query Match 9.0%; Score 170; DB 4; Length 728;
 Best Local Similarity 31.5%; Pred. No. 1.68e-05;
 Matches 28; Conservative 22; Mismatches 33; Indels 6; Gaps 4;

Db 391 Cyrgngknymgnlsqtrs gltcsmwdknmedlhrhifwepdasklnenycrnpddahgp 450
 |:||| :||| :||| | :||| :||| :||| :||| :||| :||| :||| :|||
 Qy 25 CFWDNGHLYREDQTS PAPGLRCLNW ---LDA-QSGLASAPVSGAGNHSYCRNPDEDPRGP 80

Db 451 wcy-tgnplip-wdyccpisrcgdtpti 477
 |:||| :||| :||| | :||| :||| :||| :||| :||| :|||
 Qy 81 WCYVSGEAGVPEKRPCEDLRCPETTSQAL 109

Search completed: Fri Sep 17 18:25:31 1999
 Job time : 105 secs.

The page is framed by a decorative border. The left and right sides are defined by vertical lines with a repeating pattern of small black stars. The top and bottom edges are formed by horizontal dashed lines. Inside this border, there are five sets of dashed lines arranged vertically, each set consisting of a large rectangle with internal vertical and horizontal lines forming a grid-like structure.

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Fri Sep 17 18:29:09 1999; MasPar time 4.08 Seconds
654.758 Million cell updates/sec
tabular output not generated.

>>US-09-084-491A-2
(1-263) from US09084491A.pep

Sequence: 1 MLLAWVQAFLVSNMLLAEAY.....PVDPQEGSTPLMGQQAGTPGA 263

Searched: 106580 seqs, 10152877 residues
Gap 11

Database: a-issued
Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	184.	9.8	812	3 PCT-US95-0	Sequence 1, Application	8.64e-07
2	184	9.8	812	2 US-08-612-	Sequence 1, Application	8.64e-07
3	184	9.8	812	2 US-08-605-	Sequence 1, Application	8.64e-07
4	184	9.8	812	2 US-08-452-	Sequence 1, Application	8.64e-07
5	184	9.8	812	2 US-08-429-	Sequence 1, Application	8.64e-07
6	184	9.8	812	1 US-08-451-	Sequence 1, Application	8.64e-07
7	184	9.8	812	2 US-08-326-	Sequence 1, Application	8.64e-07
8	184	9.8	812	1 US-08-248-	Sequence 1, Application	8.64e-07
9	181	9.6	655	1 US-08-148-	Sequence 12, Application	1.54e-06
10	181	9.6	655	1 US-08-448-	Sequence 12, Application	1.54e-06
11	176	9.3	217	2 US-08-622-	Sequence 24, Application	4.04e-06
12	167	8.9	723	1 US-08-290-	Sequence 3, Application	2.26e-05
13	167	8.9	723	1 US-08-290-	Sequence 2, Application	2.26e-05
14	167	8.9	723	1 US-08-290-	Sequence 1, Application	2.26e-05
15	167	8.9	723	1 US-08-404-	Sequence 1, Application	2.26e-05
16	167	8.9	723	1 US-07-838-	Sequence 1, Application	2.26e-05
17	167	8.9	728	1 US-07-815-	Sequence 2, Application	2.26e-05
18	167	8.9	728	1 US-08-087-	Sequence 22, Application	2.26e-05
19	165	8.8	579	1 US-08-295-	Sequence 4, Application	3.30e-05
20	165	8.8	579	3 PCT-US92-1	Sequence 4, Application	3.30e-05
21	165	8.8	615	1 US-07-998-	Sequence 3, Application	3.30e-05
22	165	8.8	615	3 PCT-US92-1	Sequence 3, Application	3.30e-05
23				US-08-162-	Sequence 2, Application	3.30e-05

24	US-08-463-	Sequence 3,	Applicatio	3.30e-05
25	US-08-184-	Sequence 8,	Applicatio	2.73e-05
26	PCT-US95-1	Sequence 2,	Applicatio	2.73e-05
27	US-08-334-	Sequence 2,	Applicatio	2.73e-05
28	US-08-469-	Sequence 54,	Applicati	7.04e-05
29	US-08-643-	Sequence 1,	Applicatio	7.04e-05
30	5200340-8	Patent No.	5200340.	7.04e-05
615	US-08-463-	Sequence 2,	Applicatio	7.04e-05
711	US-08-184-	Sequence 29,	Applicati	7.04e-05
711	PCT-US95-1	Sequence 1,	Applicatio	7.04e-05
161	US-08-334-	Sequence 11,	Applicati	1.24e-04
161	US-08-469-	Sequence 41,	Applicati	1.03e-04
161	US-08-643-	Sequence 35,	Applicati	1.50e-04
161	8.6	810	810	7.04e-05
161	8.6	810	810	7.04e-05
161	8.6	810	810	7.04e-05
161	8.6	814	814	7.04e-05
161	8.6	814	814	7.04e-05
158	8.4	101	2	US-08-643-
159	8.4	378	2	US-08-612-
157	8.3	160	2	US-08-612-
331	PCT-US95-0	339	3	Sequence 3,
332	US-08-429-	339	2	Sequence 3,
333	US-08-326-	339	2	Sequence 3,
334	US-08-248-	339	1	Sequence 3,
335	US-08-612-	339	2	Sequence 3,
336	US-08-451-	339	1	Sequence 3,
337	US-08-452-	339	2	Sequence 3,
338	US-08-612-	339	1	Sequence 40,
339	US-08-612-	339	2	Sequence 42,
40	US-08-612-	339	2	US-08-612-
41	US-08-452-	339	1	US-08-612-
42	US-08-612-	339	2	US-08-612-
43	US-08-612-	339	2	US-08-612-
44	US-08-612-	352	2	US-08-612-
45	US-08-612-	378	2	US-08-612-

RESULT 1
ID PCT-US95-05107-1
XX STANDARD; PRT; 812 AA.
XX

Sequence 1, Application PC/TUS9505107

GENERAL INFORMATION:

APPLICANT: THE CHILDREN'S MEDICAL CENTER, CORPORATION
TITLE OF INVENTION: Angiotatin and Method of Use
NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Jones & Askew
STREET: 191 Peachtree Street, 37th Floor
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.
ZIP: 30303-1769

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/05107
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/248,629
FILING DATE: 26-APR-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/326,785
FILING DATE: 20-OCT-1994

ATTORNEY/AGENT INFORMATION:

NAME: Johnson, James D.
REGISTRATION NUMBER: 31,771
REFERENCE/DOCKET NUMBER: 05213-0122
TELECOMMUNICATION INFORMATION:

TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 812 amino acids

CC STRANDEDNESS: Linear
 CC MOLECULE TYPE: protein
 CC HYPOTHETICAL: NO
 CC ORIGINAL SOURCE:
 CC ORGANISM: Murine
 SEQUENCE 812 AA; 90846 MW; 3411937 CN;
 SQ SEQUENCE 812 AA; 90846 MW; 3411937 CN;
 Query Match 9.8%; Score 184; DB 3; Length 812;
 Best Local Similarity 27.8%; Pred. No. 8.64e-07;
 Matches 35; Conservative 38; Mismatches 43; Indels 10; Gaps 9;
 Matches 35; Conservative 38; Mismatches 43; Indels 10; Gaps 9;
 Db 377 CYQSDGOSYRGTSSTTITGKKCQSWAAMFPHRSKTPENFPDAGLEMNYCRNPDGDD-KGP 435
 | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 436 WCYTTC-PPSQR-WEYCNLKRSETGG-SVVELPTVSQEPPS-GPSDSETDCMYGGKDYRG 491
 ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
 Db 492 KTAVTA 497
 QY 81 WCYVSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQAEGPAADEVQ-VFAPANALPA 139
 QY 140 RSEAAA 145
 QY 140 RSEAAA 145
 RESULT 2
 ID US-08-612-788-1 STANDARD; PRT; 812 AA.
 XX XXXXXX
 XX AC
 DT XX
 DE XX
 Sequence 1, Application US/08612788
 CC Sequence 1, Application US/08612788
 CC Patent No. 5837682
 CC GENERAL INFORMATION:
 CC APPLICANT: Folkman, M. Judah
 CC APPLICANT: O'Reilly, Michael
 CC APPLICANT: Cao, Yihai
 CC APPLICANT: Sim, B. Kim Lee
 CC TITLE OF INVENTION: Angiotatin Fragments and Method of Use
 CC NUMBER OF SEQUENCES: 45
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Jones & Askew
 CC STREET: 191 Peachtree Street, 37th Floor
 CC CITY: Atlanta
 CC STATE: Georgia
 CC COUNTRY: U.S.
 CC ZIP: 30303-1769
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patentin Release #1.0, Version #1.30
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/612,788
 CC FILING DATE:
 CC CLASSIFICATION: 514
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Warren, William L.
 CC REGISTRATION NUMBER: 36,714
 CC REFERENCE/DOCKET NUMBER: 05213-0126
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 404-818-3700
 CC TELEFAX: 404-818-3799
 CC INFORMATION FOR SEQ ID NO: 1:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 812 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: Linear
 CC TOPOLogy: linear
 CC MOLECULE TYPE: protein
 CC MOLECULE TYPE: peptide

CC ORIGINAL SOURCE: Murine Plasminogen
 CC ORGANISM: Murine Plasminogen
 SQ SEQUENCE 812 AA; 90846 MW; 3411937 CN;
 Query Match 9.8%; Score 184; DB 2; Length 812;
 Best Local Similarity 27.8%; Pred. No. 8.64e-07;
 Matches 35; Conservative 38; Mismatches 43; Indels 10; Gaps 9;
 Db 377 CYQSDGQSRYRGTSSTTITGKKCQSWAAMFPHRSKTPENFPDAGLEMNYCRNPDD-KGP 435
 QY ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:
 25 CFWDNGHLYREDQTPAPGLRCLNWLDS-AQS-GLASAPVSGAG-NHSYCRNPDEDPRGP 80
 Db 436 WCYT-DPSVR-WEYCNLKCSETGG-SVVELPTVSQEPS-GPSDSETDCMYGNGKDYRG 491
 QY ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:
 492 KTAVTA 497
 Db 436 WCYT-DPSVR-WEYCNLKCSETGG-SVVELPTVSQEPS-GPSDSETDCMYGNGKDYRG 491
 QY ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:
 81 WCYVSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASEGPGADEVQ-VFAPANALPA 139
 Db 492 KTAVTA 497
 QY ||:||:||:
 140 RSEAAA 145
 RESULT 5
 ID US-08-429-743-1 STANDARD; PRT; 812 AA.
 XX XXXXXXXX
 AC XXXXXX
 DT XX
 DE Sequence 1, Application US/08429743
 XX XX
 CC Sequence 1, Application US/08429743
 CC Patent No. 5885795
 CC GENERAL INFORMATION:
 CC APPLICANT: O'Reilly, Michael
 CC APPLICANT: Folkman, M. Judah
 CC APPLICANT: Sim, Kim Lee
 CC APPLICANT: Cao, Yihai
 CC TITLE OF INVENTION: Angiostatin and Method of Use
 CC NUMBER OF SEQUENCES: 6
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Jones & Askew
 CC STREET: 191 Peachtree Street, 37th Floor
 CC CITY: Atlanta
 CC STATE: Georgia
 CC COUNTRY: U.S.
 CC ZIP: 30303-1769
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn Release #1.0, Version #1.30
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/429,743
 CC FILING DATE:
 CC CLASSIFICATION: 424
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 08/248,629
 CC FILING DATE: 26-APR-1994
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 08/326,785
 CC FILING DATE: 20-OCT-1994
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Johnson, James D.
 CC REGISTRATION NUMBER: 31,771
 CC REFERENCE/DOCKET NUMBER: 05213-0122
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 404-818-3700
 CC TELEFAX: 404-818-3799
 CC INFORMATION FOR SEQ ID NO: 1:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 812
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC SEQUENCE 812 AA; 90846 MW; 3411937 CN;
 SQ Query Match 9.8%; Score 184; DB 2; Length 812;
 Best Local Similarity 27.8%; Pred. No. 8.64e-07;

Query Match 9.8%; Score 184; DB 2; Length 812;
 Best Local Similarity 27.8%; Pred. No. 8.64e-07;
 Matches 35; Conservative 38; Mismatches 43; Indels 10; Gaps 9;

Db 377 CYQSDGOSYRGTSSTTITGKKCQSWAAMFPHRHSKTPENFPDAGLEMNYCRNPNDGD-KGP 435
 Qy 25 CFWDNGHLYREDQTSPAPGLRCLNWLD--AQ5-GLASAPVSGAG-NHSYCRNPDEDPRG 80

Db 436 WCYT-DPSVR-WEYCNLKRCSETGG-SVVELPTVSQEPS-GPSDSETDCMYGNGKDYRG 491
 Qy 81 WCYVSGEAGVPEKRPCETTSQALPAFTTEIQEASEGPGADEVQ-VFAPANALPA 139

Db 436 WCYT-DPSVR-WEYCNLKRCSETGG-SVVELPTVSQEPS-GPSDSETDCMYGNGKDYRG 491
 Qy 81 WCYVSGEAGVPEKRPCETTSQALPAFTTEIQEASEGPGADEVQ-VFAPANALPA 139

Db 492 KTAVTA 497
 Qy 140 RSEAAA 145

RESULT 7
 ID US-08-326-785-1 STANDARD; PRT; 812 AA.
 XX
 AC XXXXX
 XX
 DT
 XX
 DE Sequence 1, Application US/08326785
 XX
 CC Sequence 1, Application US/08326785
 CC Patent No. 5792845

GENERAL INFORMATION:
 APPLICANT: Folkman, Judah
 APPLICANT: O'Reilly, Michael
 TITLE OF INVENTION: Angiostatin and Method of Use
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Jones & Askew
 STREET: 191 Peachtree Street, 37th Floor
 CITY: Atlanta
 STATE: Georgia
 COUNTRY: USA
 ZIP: 30303-1769

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50
 COMPUTER: Macintosh
 OPERATING SYSTEM: 7.0
 COMPUTER: Macintosh
 OPERATING SYSTEM: 7.0
 SOFTWARE: Microsoft Word
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/326,785
 FILING DATE:
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/248,629

ATTORNEY/AGENT INFORMATION:
 NAME: Larry W. Stults, Ph.D.
 REGISTRATION NUMBER: 34,025
 REFERENCE/DOCKET NUMBER: 05213-0121
 FILING DATE: 04/26/94
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 404-818-3700
 TELEFAX: 404-818-3799
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 812
 TYPE: amino acid
 TOPOLOGY: linear
 SEQUENCE 812 AA; 90846 MW; 3411937 CN;

CC SQ Query Match 9.8%; Score 184; DB 2; Length 812;
 Best Local Similarity 27.8%; Pred. No. 8.64e-07;
 Matches 35; Conservative 38; Mismatches 43; Indels 10; Gaps 9

Db 377 CYOSDGOSYRGTSSTTITGKKCQSWAAMFPHRHSKTPENFPDAGLEMNYCRNPNDGD-KGP 435
 Qy 25 CFWDNGHLYREDQTSPAPGLRCLNWLD--AQ5-GLASAPVSGAG-NHSYCRNPDEDPRG 80

Db 436 WCYT-DPSVR-WEYCNLKRCSETGG-SVVELPTVSQEPS-GPSDSETDCMYGNGKDYRG 491
 Qy 81 WCYVSGEAGVPEKRPCETTSQALPAFTTEIQEASEGPGADEVQ-VFAPANALPA 139

Query Match 9.8%; Score 184; DB 1; Length 812;
 Best Local Similarity 27.8%; Pred. No. 8.64e-07;
 Matches 35; Conservative 38; Mismatches 43; Indels 10; Gaps 9;

CC SQ Query Match 9.8%; Score 184; DB 1; Length 812;
 Best Local Similarity 27.8%; Pred. No. 8.64e-07;
 Matches 35; Conservative 38; Mismatches 43; Indels 10; Gaps 9

Db 377 CYOSDGOSYRGTSSTTITGKKCQSWAAMFPHRHSKTPENFPDAGLEMNYCRNPNDGD-KGP 435
 Qy 25 CFWDNGHLYREDQTSPAPGLRCLNWLD--AQ5-GLASAPVSGAG-NHSYCRNPDEDPRG 80

Db 436 WCYT-DPSVR-WEYCNLKRCSETGG-SVVELPTVSQEPS-GPSDSETDCMYGNGKDYRG 491
 Qy 81 WCYVSGEAGVPEKRPCETTSQALPAFTTEIQEASEGPGADEVQ-VFAPANALPA 139

RESULT 8
 ID US-08-248-629A-1 STANDARD; PRT; 812 AA.
 XX :|:
 QY 140 RSEAAA 145

Sequence 1, Application US/08248629A
 Patent No. 5639725
 GENERAL INFORMATION:
 APPLICANT: Folkmann, Judah
 ADDRESS: Jones & Askew
 STREET: 191 Peachtree Street, 37th Floor
 CITY: Atlanta
 STATE: Georgia
 COUNTRY: USA
 ZIP: 30303-1769
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50
 COMPUTER: Macintosh
 OPERATING SYSTEM: 7.0
 SOFTWARE: Microsoft Word
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/248,629A
 FILING DATE: 04/26/94
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Larry W. Stults, Ph.D.
 REGISTRATION NUMBER: 34,025
 REFERENCE/DOCKET NUMBER: 05213-0120
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 404-818-3700
 TELEFAX: 404-818-3799
 INFORMATION FOR SEQ ID NO: 1:
 LENGTH: 812
 TYPE: amino acid
 TOPOLOGY: linear
 SEQUENCE: 812 AA; 90846 MW; 3411937 CN;

Query Match 9.8%; Score 184; DB 1; Length 812;
 Best Local Similarity 27.8%; Pred. No. 8.64e-07;
 Matches 35; Conservative 38; Mismatches 43; Indels 10; Gaps 9;

Db 377 CYQSDGQSYYRGTSSTTITGKKCQSWAAMFPHRHSKTPENFPDAGLEMNYCRNPDD-KGP 435
 |: :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Qy 25 CFWDNGHLYREDQTSPAPGLRCLNWLD--AQS-GLASAPVSGAG-NHSYCRNPDEDPRGP 80

Db 436 WCYTT-DPSVR-WEYCNLKRCSSETGG-SVVELPTVSQEPPS-GPSDSETDCMYGNGKDYG 491
 ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Qy 81 WCYVSGEAVPEKRPCEDLRCPETSQLPAFTTEIQEASEGPGADEVQ-VFAPANALPA 139

Db 492 KTAVTA 497
 ::|:
 QY 140 RSEAAA 145

RESULT 9
 ID US-08-148-910-12 STANDARD; PRT; 655 AA.
 XX :|:
 QY 140 RSEAAA 145

DE Sequence 12, Application US/08148910
 XX
 CC Sequence 12, Application US/08148910
 CC
 CC GENERAL INFORMATION:
 CC APPLICANT: Takeshi SHIMOMURA et al.
 CC TITLE OF INVENTION: NO. 5466593el Protein and Gene Encoding Said Protein
 CC Patent No. 5466593
 CC
 CC NUMBER OF SEQUENCES: 14
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Wenderoth, Lind & Ponack
 CC STREET: 805 Fifteenth Street, N.W., #700
 CC CITY: Washington
 CC STATE: D.C.
 CC COUNTRY: U.S.A.
 CC ZIP: 20005
 COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Diskette, 5.25 inch,
 CC COMPUTER: 500 Kb Storage
 CC COMPUTER: IBM Compatible
 CC OPERATING SYSTEM: MS-DOS
 CC SOFTWARE: Wordperfect
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/148,910
 CC FILING DATE: NO. 5466593ember 5, 1993
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER:
 CC FILING DATE:
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Warren M. Cheek, Jr.
 CC REGISTRATION NUMBER: 33,367
 CC REFERENCE/DOCKET NUMBER:
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 202-371-8850
 CC TELEFAX: 202-371-8856
 CC TELEX:
 CC INFORMATION FOR SEQ ID NO: 12:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 655 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC ORIGINAL SOURCE:
 CC ORGANISM: human
 SQ SEQUENCE 655 AA; 70681 MW; 2184486 CN;

Query Match 9.6%; Score 181; DB 1; Length 655;
 Best Local Similarity 40.5%; Pred. No. 1.54e-06;
 Matches 30; Conservative 19; Mismatches 17; Indels 8; Gaps 4;

Db 286 CFLGNGTGYRGVASTSAGLSCLAWNSDLYQELHVDSVGAALLGLGPHAYCRNPDNDE 345
 ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Qy 25 CFWDNGHLYREDQTSPAPGLRCLNW---L--DAQ-SGLASAPVSGAGNHSYCRNPDED 77

Db 346 R-PWCYVVKDSALS 358
 ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
 Qy 78 RGPWCYVSGEAGVP 91

RESULT 10
 ID US-08-448-937A-12 STANDARD; PRT; 655 AA.
 XX
 AC
 XX
 DT
 DE Sequence 12, Application US/08448937A

XX Sequence 12, Application US/08448937A
 CC Patent No. 5677164
 CC GENERAL INFORMATION:
 CC APPLICANT: Takeshi SHIMOMURA et al.
 CC TITLE OF INVENTION: No. 5677164el Protein and Gene Encoding Said Protein
 CC NUMBER OF SEQUENCES: 14
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Wenderoth, Lind & Ponack
 CC STREET: 805 Fifteenth Street, N.W., #700
 CC CITY: Washington
 CC STATE: D.C.
 CC COUNTRY: U.S.A.
 CC ZIP: 20005
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Diskette, 5.25 inch,
 CC COMPUTER: IBM Compatible
 CC OPERATING SYSTEM: MS-DOS
 CC SOFTWARE: Wordperfect
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/448,937A
 CC FILING DATE: May 24, 1995
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/148,910
 CC FILING DATE: NO. 5677164ember 5, 1993
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Warren M. Cheek, Jr.
 CC REGISTRATION NUMBER: 33,367
 CC REFERENCE/DOCKET NUMBER:
 CC TELECOMMUNICATION INFORMATION:
 CC - TELEPHONE: 202-371-8850
 CC - TELEFAX: 202-371-8856
 CC TELEX:
 CC INFORMATION FOR SEQ ID NO: 12:
 CC - SEQUENCE CHARACTERISTICS:
 CC LENGTH: 655 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC ORIGINAL SOURCE:
 CC ORGANISM: human
 SQ SEQUENCE 655 AA; 70681 MW; 2184486 CN;

Query Match 9.6%; Score 181; DB 1; Length 655;
 Best Local Similarity 40.5%; Pred. No. 1.54e-06;
 Matches 30; Conservative 19; Mismatches 17; Indels 8; Gaps 4;

Db 286 CFLGNGTGYRGVASTSASGLSCLAWNSDLLYQELHVDSVGAALLGLPHAYCRNPNDNE 345
 QY 25 CFWDNGHLYREDQTSPAPGLRCLNW---L--DAQ-SGLASAPVSGAGNHSYCRNPDEDP 77

Db 346 R-PWCVVVKDSALS 358
 QY 78 RGWPWCYVSGEAGVP 91

RESULT 11 ID US-08-622-720A-24 STANDARD; PRT; 217 AA.
 XX AC xxxxxxxx
 DT
 XX
 DE Sequence 3, Application US/08290937B
 QY 25 CFWDNGHLYREDQTSPAPGLRCLNW-LDA-Q-SGLASAPVSGAG-NHSYCRNPDEDP 80

Db 87 CYHAGAEYLRYGSVSKTRKGIRCONWSAETPHKPQFKHTSAPHTPLEENFCRNPDDSHGP 146
 QY 25 CFWDNGHLYREDQTSPAPGLRCLNW-LDA-Q-SGLASAPVSGAG-NHSYCRNPDEDP 80

Db 147 WCYTT-DPGTP 156
 QY 81 WCYVSGEAGVP 91

RESULT 12 ID US-08-290-937B-3 STANDARD; PRT; 723 AA.
 XX AC xxxxxxxx
 DT
 XX
 DE Sequence 3, Application US/08290937B
 QY 25 CFWDNGHLYREDQTSPAPGLRCLNW-LDA-Q-SGLASAPVSGAG-NHSYCRNPDEDP 80

Db 5648233 GENERAL INFORMATION:
 APPLICANT: YAMAGUCHI, KYOJI
 APPLICANT: SHIMA, NOBUYUKI
 APPLICANT: MURAKAMI, AKIHIKO
 APPLICANT: GOTO, MASAAKI
 APPLICANT: TSUDA, EISUKE
 APPLICANT: MASUNAGA, HIROAKI
 APPLICANT: TAKAHIRA, REIKO
 APPLICANT: OOGAKI, FUMIKO
 APPLICANT: UEDA, MASATSUGU
 APPLICANT: HIGASHIO, KANJI
 TITLE OF INVENTION: MODIFIED TCF
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Testa, Burwitz & Thibeault
 STREET: 125 High St.
 CITY: Boston
 STATE: MA
 COUNTRY: USA

Sequence 24, Application US/08622720A
 CC Patent No. 5814308
 CC GENERAL INFORMATION:
 CC APPLICANT: Zhang, Ke
 CC TITLE OF INVENTION: METHODS FOR THE TREATMENT OF

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/290,937B
 FILING DATE: 19-AUG-1994
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: CAMPBELL, PAULA A.
 REGISTRATION NUMBER: 32,503
 REFERENCE/DOCKET NUMBER: FJN-022
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 248-7000
 TELEFAX: (617) 248-7100
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 723 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE 723 AA; 82586 MW; 2690825 CN;

 Query Match 8.98; Score 167; DB 1; Length
 Best Local Similarity 31.5%; Pred. No. 2.26e-05;
 Matches 28; Conservative 22; Mismatches 33; Inde-

 Db 386 CYRGNNGKNYMGNLSQLTRSGLTCMSMWDKNMEDLHRHIFWEPDASKLNENY
 |::|||:|:|:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 QY 25 CFWDNGHLYREDQTSAPGLRCLNW--LDA-QSGLASAPVSGAGNHSY

 Db 446 WCY-TGNPLIP-WDYCPISRCEGDTTPTI 472
 |||:|::|:|:|||:|||:|||:|||:|||:
 QY 81 WCYVSGEAGVPEKRPCEDLRCPETTSQAL 109

 RESULT 14
 ID US-08-290-937B-1 STANDARD; PRT; 723 AA.
 XX XXXXXX
 AC XX
 DT XX
 DE Sequence 1, Application US/08290937B
 CC Patent No. 5648233
 CC GENERAL INFORMATION:
 CC APPLICANT: YAMAGUCHI, KYOJI
 CC APPLICANT: SHIMA, NOBUYUKI
 CC APPLICANT: MURAKAMI, AKIHIKO
 CC APPLICANT: GOTO, MASAAKI
 CC APPLICANT: TSUDA, EISUKE
 CC APPLICANT: MASUNAGA, HIROAKI
 CC APPLICANT: TAKAHIRA, REIKO
 CC APPLICANT: OOGAKI, FUMIKO
 CC APPLICANT: UEDA, MASATSUGU
 CC APPLICANT: HIGASHIO, KANJI
 CC TITLE OF INVENTION: MODIFIED TCF
 CC NUMBER OF SEQUENCES: 13
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Testa, Hurwitz & Thibeault
 CC STREET: 125 High St.
 CC CITY: Boston
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02110
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn Release #1.0, Version #1.30
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/290,937B

CC FILING DATE: 19-AUG-1994
 CC CLASSIFICATION: 530
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: CAMPBELL, PAULA A.
 CC REGISTRATION NUMBER: 32, 503
 CC REFERENCE/DOCKET NUMBER: FJN-022
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 248-7000
 CC TELEFAX: (617) 248-7100
 CC INFORMATION FOR SEQ ID NO: 1:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 723 amino acids
 CC STRANDEDNESS: single
 CC TOPOLogy: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE 723 AA; 82602 MW; 2699270 CN;
 CC
 Query Match 8.9%; Score 167; DB 1; Length 723;
 Best Local Similarity 31.5%; Pred. No. 2.26e-05;
 Matches 28; Conservative 22; Mismatches 33; Indels 6; Gaps 4;
 Db 386 CYRGNNGKNYMGNLSQLTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCRNPDDDAHGP 445
 Qy 25 CFWDNGHLYREDQTSPAPGLRCLNW--LDA-QSGLASAPVSGAGNHSYCRNPDDEDPRGP 80
 Sq 446 WCY-TGNPLIP-WDYCPISRCEGDTPTI 472
 Qy 81 WCYVSGEAGVPEKRPCEDLRCPETTSQAL 109
 Search completed: Fri Sep 17 18:29:44 1999
 Job time : 35 secs.

Query Match 8.9%; Score 167; DB 1; Length 723;
 Best Local Similarity 31.5%; Pred. No. 2.26e-05;
 Matches 28; Conservative 22; Mismatches 33; Indels 6; Gaps 4;
 Db 386 CYRGNNGKNYMGNLSQLTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNENYCRNPDDDAHGP 445
 Qy 25 CFWDNGHLYREDQTSPAPGLRCLNW--LDA-QSGLASAPVSGAGNHSYCRNPDDEDPRGP 80
 Db 446 WCY-TGNPLIP-WDYCPISRCEGDTPTI 472
 Qy 81 WCYVSGEAGVPEKRPCEDLRCPETTSQAL 109

RESULT 15
 ID JS-08-404-643-1 STANDARD; PRT; 723 AA.
 XX XXXXXX
 AC XX
 DT XX
 DE XX
 Sequence 1, Application US/08404643
 CC Sequence 1, Application US/08404643
 CC Patent No. 5658742
 CC GENERAL INFORMATION:
 CC APPLICANT: HIGASHIO, KANJI
 CC APPLICANT: SHIMA, NOBUYUKI
 CC APPLICANT: OOGAKI, FUMIKO
 CC TITLE OF INVENTION: MONOCLONAL ANTIBODY
 CC NUMBER OF SEQUENCES: 1
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
 CC ADDRESSEE: THIBEAULT
 CC STREET: 53 STATE STREET
 CC CITY: BOSTON
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02109
 COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent Release #1.0, version #1.25
 CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/404, 643
 CC FILING DATE: 15-MAR-1995
 CC CLASSIFICATION: 436
 ATTORNEY/AGENT INFORMATION:
 CC NAME: CAMPBELL, PAULA A
 CC REGISTRATION NUMBER: 32, 503
 CC REFERENCE/DOCKET NUMBER: FJN-034 (3999/35)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617)-248-7000
 CC INFORMATION FOR SEQ ID NO: 1:

RESULT 2 #type complete
 ENTRY PLMS #type complete
 plasmin (EC 3.4.21.7) precursor - mouse
 CONTAINS angiotatin; Plasminogen
 ORGANISM #formal_name Mus musculus #common_name house mouse
 DATE 20-Sep-1991 #sequence_revision 01-Nov-1996 #text_change
 08-May-1998
 ACESSONS A38514; S48202; S48203
 REFERENCE #authors Degen, S.J.F.; Bell, S.M.; Schaefer, L.A.; Elliott, R.W.
 #journal Genomics (1990) 8:49-61
 #title Characterization of the cDNA coding for mouse plasminogen and
 localization of the gene to mouse chromosome 17.
 #cross-references MUID:91184812
 #accession A38514
 ##molecule_type mRNA
 ##residues 1-812 #label DEG
 ##cross-references GB:J04766; NID:9200402; PID:g200403
 REFERENCE S48202
 #authors Linnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
 #journal Eur. J. Biochem. (1994) 224:863-871
 #title Characterization of the murine plasma fibrinolytic system.
 #cross-references MUID:95010076
 #accession S48202
 ##molecule_type protein
 ##residues 20-25 #label LIJ
 #accession S48203
 ##molecule_type protein
 ##residues 22-27 #label LI2
 COMMENT Plasminogen is synthesized by the kidney and is present in plasma
 and many other extracellular fluids.
 Plasminogen is converted into plasmin by plasminogen activators,
 both plasminogen and its activator being bound to fibrin. Plasmin
 is inactivated by alpha-2-antiplasmin (see PIR:S47217)
 immediately after dissociation from the clot. In the presence of
 the inhibitor, the activation involves only cleavage after
 Arg-581, resulting in two chains connected by two disulfide
 bonds. Without the inhibitor, the activation involves also
 removal of the activation peptide.

COMMENT Stromelysin 1 (see PIR:KCMSS1) acts on plasminogen to produce
 angiotatin. Together with endostatin (see PIR:A56101,
 PIR:B56101), angiostatin acts to inhibit angiogenesis, and so may
 be useful in treating solid tumors.

FUNCTION #description dissolves the fibrin of blood clots; acts as a proteolytic
 factor in a variety of processes including embryonic
 development, tissue remodeling and tumor invasion; in
 ovulation it weakens the walls of the graafian follicle;
 also activates the urokinase-type plasminogen activator
 fibrinolysis

#pathway #superfamily plasmin; kringle homology; plasminogen-related
 protein precursor homology; trypsin homology
 angiogenesis inhibitor; blood; duplication; fibrinolysis;
 glycoprotein; hydrolase; kidney; kringle; plasma; serine
 proteinase; zymogen

CLASSIFICATION
 KEYWORDS

FEATURE 1-96
 #domain plasminogen-related protein precursor homology
 #label PLPH\\
 #domain signal sequence #status Predicted #label SIG\\
 #product plasminogen #status Predicted #label PRO\\
 #domain activation peptide #status Predicted #label APT\\
 #product angiostatin #status Predicted #label AST\\
 #product plasmin #status Predicted #label MAT\\
 #domain chain A #status Predicted #label MAT\\
 #domain kringle homology #label K1\\
 #domain kringle homology #label KR1\\
 #domain kringle homology #label KR2\\
 #domain kringle homology #label KR3\\
 #domain kringle homology #label KR4\\
 #domain kringle homology #label KR5\\
 #domain chain B #status Predicted #label BCH\\
 #domain trypsin homology #label TRY\\
 1-19
 20-812
 20-96
 79-466
 97-581, 582-812
 103-181
 185-262
 275-352
 377-454
 481-560
 582-812
 582-805
 49-73, 53-61,
 103-181, 124-164,

152-176, 185-262,
 188-316, 206-245,
 234-257, 275-352,
 296-335, 324-347,
 377-454, 398-437,
 426-449, 481-560,
 502-543, 531-555,
 568-687, 578-586,
 609-625, 701-768,
 731-747, 758-786
 78-79 #cleavage_site Glu-Asn (stromelysin 1) #status
 predicted\\
 #binding_site carbohydrate (Asn) (covalent) #status
 predicted\\
 #disulfide_bonds #status Predicted\\
 #cleavage_site stromelysin 1) #status
 predicted\\
 #binding_site carbohydiate (Asn) (covalent) #status
 predicted\\
 #cleavage_site Thr-Val (stromelysin 1) #status
 predicted\\
 #cleavage_site Arg-Val (plasminogen activator) #status
 experimental\\
 active_site His, ASP, Ser #status Predicted
 SUMMARY 9.8%; Score 184; DB 1; Length 812;
 Best Local Similarity 27.8%; Pred. No. 7.26e-14;
 Matches 38; Mismatches 43; Indels 10; Gaps 9;
 Db 377 CYQSDGQSYRGTSSTTITGKKCQSWAAMFPHRHSKTPENFDAGLEMNYCRNPDGDKGP 435
 |:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:
 Qy 25 CFWDNGHLYREDQTSPAPGLRCLNWLD--AQG-GLASAPVSGAG-NHSYCRNPDEDPRGP 80
 Db 436 WCYTT-DPSYR-WEYCNLKRCSETGG-SVVELPTVSQEPS-GPSDSETDCMYGNGKDYRG 491
 |||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:
 Qy 81 WCYSGEAGVPKEPKPCEDLRCPETTSQALPAFTTEIQEASEGPGADEVQ-VFAPPNALPA 139
 Db 492 KTAVTA 497
 ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:
 Qy 140 RSEAA 145
 RESULT 3 #type complete
 ENTRY S00657 #type complete
 apoprotein(a) (EC 3.4.21.-) precursor - human
 ALTERNATE_NAMES apolipoprotein(a) chain apo(a)
 ORGANISM #formal_name Homo sapiens #common_name man
 DATE 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change
 ACCESSIONS S00657; A28017; A47277; I60906; A47233; I52415; I65286
 REFERENCE S00657
 #authors McLean, J.W.; Tomlinson, J.E.; Kuang, W.J.; Eaton, D.L.;
 Chen, E.Y.; Fless, G.M.; Scanu, A.M.; Lawn, R.M.
 #journal Nature (1987) 330:132-137
 #title CDNA sequence of human apolipoprotein(a) is homologous to
 plasminogen.
 #cross-references MUID:88039109
 #accession S00657
 ##molecule_type mRNA
 #residues 1-4548 #label MCL
 ##cross-references GB:X06290; EMBL:X06696; NID:g28619; PID:g28620
 REFERENCE A28017
 #authors Eaton, D.L.; Fless, G.M.; Kohr, W.J.; McLean, J.W.; Xu, Q.T.;
 Miller, C.G.; Lawn, R.M.; Scanu, A.M.
 #journal Proc. Natl. Acad. Sci. U.S.A. (1987) 84:3224-3228
 #title Partial amino acid sequence of apolipoprotein(a) shows that
 it is homologous to plasminogen.
 #cross-references MUID:87204109
 #accession A28017
 ##molecule_type protein
 ##residues 20-21, 'P', 23-34, 177-179, 'N', 181-186, 'T', 188-196, 'DKG',
 200-292-314, 'W', 316-318, 4201-4202, 'X', 4204-4207, 'LL',
 4210-4212, 'P', 4214-4218, 'X', 4220-4232, 4323-4334, 'LTP',
 4382-4386, 'X', 4388-4393, 'PX', 4396-4401 #label EAT
 REFERENCE A47277
 #authors Wade, D.P.; Clarke, J.G.; Lindahl, G.E.; Liu, A.C.; Zysow,
 B.R.; Meer, K.; Schwartz, K.; Lawn, R.M.


```

#authors Coffer, A.; Fellows, J.; Young, S.; Pappin, D.; Rahman, D. RESULT 13
#journal Biochem. J. (1991) 278:35-41 JH0579 #type complete
#title Purification and characterization of biologically active hepatocyte growth factor precursor - human scatter factor from ras-transformed NIH 3T3 conditioned medium. ALTERNATE_NAMES
#cross-references MUID:91354223 ORGANISM
#accession S17173 #formal_name Homo sapiens #common_name man DATE 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change
18-Sep-1998
#molecule_type protein JH0579; JU0333; A41140; B36677; A33512; A39006;
#residues 496-517, 'T', 519 ##label COF PH0114; A37794; I59214; S15443
REFERENCE JH0579 #authors Seki, T.; Hagiya, M.; Shimonishi, M.; Nakamura, T.; Shimizu, S.
#journal Gene (1991) 102:213-219
#title Organization of the human hepatocyte growth factor-encoding gene. #cross-references MUID:91340155
#accession JH0579
#molecule_type protein #molecule_type DNA
#residues 1-728 ##label SEK
#cross-references DDBJ:D90318
#note the authors translated the codon GAA for residue 662 as Gly
REFERENCE JU0333 #authors Seki, T.; Hagiya, M.; Shimonishi, M.; Nakamura, T.; Shimizu, S.
#submission submitted to JIPID, March 1991
#description Organization of the human hepatocyte growth factor-encoding gene.
#accession JU0333
#molecule_type DNA
#residues 1-481, 'RT', 484-728 ##label SE2
REFERENCE A41140 #authors Weidner, K.M.; Arakaki, N.; Hartmann, G.; Vandekerckhove, J.; Weingart, S.; Rieder, H.; Fonatsch, C.; Tsubouchi, H.; Hishida, T.; Daikuhara, Y.; Birchmeier, W.
#cross-references EMBL:X81630; PID:g673451; PID:g673452 #note PRO. Natl. Acad. Sci. U.S.A. (1991) 88:7001-7005 Evidence for the identity of human scatter factor and human hepatocyte growth factor.
#cross-references MUID:91334393 #accession A41140
#molecule_type mRNA
#residues 1-728 ##label WEI
#cross-references GB:M73239; PID:g337935; PID:g337936
REFERENCE A36677 #authors Seki, T.; Ihara, I.; Sugimura, A.; Shimonishi, M.; Nishizawa, T.; Asami, O.; Hagiya, M.; Nakamura, T.; Shimizu, S.
#journal Biochem. Biophys. Res. Commun. (1990) 172:321-327
#title Isolation and expression of cDNA for different forms of hepatocyte growth factor from human leukocyte.
#cross-references MUID:91025062 #accession B36677
#molecule_type mRNA
#residues 1-728 ##label SE1
#cross-references GB:M60718; PID:g184031; PID:g184032 #accession A36677
#molecule_type mRNA
#residues 1-161,167-728 ##label SE4
#cross-references EMBL:X16323 #experimental_source leukocyte
REFERENCE A33512 #authors Miyazawa, K.; Tsubouchi, H.; Naka, D.; Takahashi, K.; Okigaki, M.; Arakaki, N.; Nakayama, H.; Hirono, S.; Sakiyama, O.; Takahashi, K.; Gohda, E.; Daikuhara, Y.; Kitamura, N.
#journal Biochem. Biophys. Res. Commun. (1989) 163:967-973
#title Molecular cloning and sequence analysis of cDNA for human hepatocyte growth factor.
#cross-references MUID:89392017 #accession A33512
#status not compared with conceptual translation
#molecule_type mRNA
#residues 1-728 ##label MIY
#cross-references GB:M29145; PID:g184041; PID:g306846
#cross-references GB:M29145; PID:g184041; PID:g306846

```

```

REFERENCE
#authors Rubin, J.S.; Chan, A.M.L.; Bottaro, D.P.; Burgess, W.H. ;
  Taylor, W.G.; Cech, A.C.; Hirschfield, D.W.; Wong, J.J. ;
  Miki, T.; Finch, P.W.; Aaronson, S.A.
#journal Proc. Natl. Acad. Sci. U.S.A. (1991) 88:415-419
#title A broad-spectrum human lung fibroblast-derived mitogen is a
  variant of hepatocyte growth factor.
#cross-references MUID:91110540
#accession A39006
##molecule_type mRNA
##residues 1-161,167-728 ##label RUB
##cross-references GB:M55379
##experimental_source embryonic lung
REFERENCE
#authors Yoshiyama, Y.; Arakaki, N.; Naka, D.; Takahashi, K.; Hirono,
  S.; Kondo, J.; Nakayama, H.; Gohda, E.; Kitamura, N. ;
  Tsubouchi, H.; Ishii, T.; Hishida, T.; Daikuhara, Y.
#journal Biochem. Biophys. Res. Commun. (1991) 175:660-667
#title Identification of the N-terminal residue of the heavy chain
  of both native and recombinant human hepatocyte growth
  factor. MUID:91207365
#accession PH0114
##molecule_type protein
##residues 32-43;53-58 ##label YOS
##experimental_source plasma
REFERENCE
#authors Weidner, K.M.; Behrens, J.; Vandekerckhove, J.; Birchmeier,
  W.
#journal J. Cell Biol. (1990) 111:2097-2108
#title Scatter factor: molecular characteristics and effect on the
  invasiveness of epithelial cells.
#cross-references MUID:91035621
#accession A37796
##molecule_type protein
##residues 86-91;329-344;356-363,'XX',366-370;425-434;442-447,'X',
  449-450;543-546,'X',548-553;563-565,'X',567-574
##label WE2
REFERENCE
#authors Nakamura, T.; Nishizawa, T.; Hagiya, M.; Seki, T. ;
  Shimonishi, M.; Sugimura, A.; Tashiro, K.; Shimizu, S.
#journal Nature (1989) 342:440-443
#title Molecular cloning and expression of human hepatocyte growth
  factor. MUID:90066676
#accession S06794
##molecule_type mRNA
##residues 1-31,'HK',34-77,'N',79-292,'V',294-299,'M',301-316,'A',
  318-335,'K',337-386,'N',388-415,'N',417-504,'V',
  506-508,'I',510-557,'E',559-560,'R',562-594,'N',
  596-728 ##label NAK
##cross-references EMBL:X16323; NID:932081; PID:g32082
##experimental_source liver
##note the authors translated the codon CAG for residue 727 as
  Glu
##note part of this sequence, including the amino end of both
  the alpha and beta chains, was confirmed by protein
  sequencing
REFERENCE
#authors Hartmann, G.; Naldini, L.; Weidner, K.M.; Vigna,
  E.; Comoglio, P.M.; Birchmeier, W.
#journal Proc. Natl. Acad. Sci. U.S.A. (1992) 89:11574-11578
#title A functional domain in the heavy chain of scatter
  factor/hepatocyte growth factor binds the c-Met receptor
  and induces cell dissociation but not mitogenesis.
#cross-references MUID:93087571
#accession I59214
##status preliminary; translated from GB/EMBL/DDBJ
##molecule_type mRNA
##residues 1-288,'ET' ##label RES
##cross-references GB:L02931; NID:g184033; PID:g184034
REFERENCE
#authors Miyazawa, K.; Kitamura, A.; Nakka, D.; Kitamura, N.

```

```

REFERENCE A00914 #authors Degen, S.J.F.; MacGillivray, R.T.A.; Davie, E.W.
#journal Biochemistry (1983) 22:2087-2097
#title Characterization of the complementary deoxyribonucleic acid
and gene coding for human prothrombin.
#cross-references MUID:83231469

#accession A00914
##molecule_type mRNA
##residues 8-163, 'N', 165-622 ##label DE2
##cross-references GB:V00595; GB:J00307; NID:g37128; PID:e5121;
PID:91335344

#accession B00914
##molecule_type DNA
##residues 188-311 ##label DE3
REFERENCE A37549
#authors Walz, D.A.; Hewett-Emmett, D.; Seegers, W.H.
#journal Proc. Natl. Acad. Sci. U.S.A. (1977) 74:1969-1972
#cross-references MUID:77193964

#accession A37549
##molecule_type protein
##residues 44-118, 'N', 120, 'S', 122-163, 'I', 165-175, 'A', 177-182, 'T',
184-193, 'M', 196-308, 'EE', 309-314 ##label WAL
REFERENCE A37550
#authors Butkowski, R.J.; Elion, J.; Downing, M.R.; Mann, K.G.
#journal J. Biol. Chem. (1977) 252:4942-4957
#title Primary structure of human prethrombin 2 and alpha-thrombin.
#cross-references MUID:77207112

#accession A37550
##molecule_type protein
##residues 315-334, 'N', 336-348, 'N', 350-368, 'N', 370-397, 'N', 399-413,
'N', 415-484, 'N', 486-493, 'G', 495-503, 'Y', 505-508, 'S',
510, 'V', 512-513, 'D', 515-528, 'A', 531, 'Q', 533-622
##label BUT

REFERENCE A37551
#authors Rabiet, M.J.; Blashill, A.; Furie, B.; Furie, B.C.
#journal J. Biol. Chem. (1986) 261:13210-13215
#cross-references MUID:87008532

#contents annotation; activation cleavages
REFERENCE I51952
#authors MacGillivray, R.T.; Irwin, D.M.; Guinto, E.R.; Stone, J.C.
#journal Ann. N. Y. Acad. Sci. (1986) 485:73-79
#title Recombinant genetic approaches to functional mapping of
thrombin.
#cross-references MUID:87182874
#accession I51952
##status translated from GB/EMBL/DDBBJ
##molecule_type mRNA
##residues 1-2, 'R', 5-100 ##label RES
##cross-references GB:M33031; NID:g190723; PID:g190724

COMMENT Thrombin, which cleaves bonds after Arg and Lys, converts
fibrinogen to fibrin and activates factors V, VIII, XIII, and, in
complex with thrombomodulin, protein C.
Prothrombin is activated on the surface of a phospholipid membrane
that binds the amino end of prothrombin and factors Va and Xa in
calcium-dependent interactions. The activation peptide(s) can be
removed either by factor Xa or thrombin; the cleavage into light
and heavy chains is by factor Xa. It is not known whether one or
two smaller activation peptides, with additional cleavage after
314-Arg, are released in natural blood clotting.
The cleavage after Arg-198, observed in vitro, does not occur in
plasma.
The gamma-carboxylglutamyl residues bind calcium ions, result from
the carboxylation of glutamyl residues by microsomal vitamin
K-dependent carboxylase, and are necessary for calcium-dependent
interaction with the negatively charged phospholipid membrane
surface.
COMMENT The prothrombin precursor is synthesized in the liver.

GENETICS
#gene GDB:F2
##cross-references GDB:119894; OMIM:176930
##map_position 11p11-11q12
##introns 27/1; 80/3; 89/1; 106/1; 141/2; 187/1; 292/1; 335/1; 377/2;

COMMENT The prothrombin precursor is synthesized in the liver.

CLASSIFICATION
#superfamily thrombin; Gla domain homology; kringle homology
#trypsin homology
#acute phase; blood coagulation; calcium binding;
#carboxyglutamic acid; duplication; glycoprotein; hydrolase;
#kringle; liver; plasma; serine proteinase

KEYWORDS
#domain signal sequence #status predicted #label SIG\
#domain kringle homology #status predicted #label PRO\
#domain kringle homology #label GLA\
#product prothrombin #status experimental #label MAT\
#domain activation peptide #status experimental #label
APT \
#domain kringle homology #label KRI\
#domain kringle homology #label KR2\
#product thrombin light chain #status experimental
#label LCH\
#product thrombin heavy chain #status experimental
#label HCH\
#domain trypsin homology #label TRY\
#modified_site gamma-carboxyglutamic acid (Glu) #status
experimental \
#modified_bonds #status predicted \
#binding_site carbohydrate (Asn) (covalent) #status
predicted \
#length 622 #molecular-weight 70036 #checksum 3003
SUMMARY
Query Match 8.8%; Score 165; DB 1; Length 622;
Best Local Similarity 37.7%; Pred. No. 1.1le-10;
Matches 26; Conservative 13; Mismatches 26; Indels 4; Gaps 3;
Db 213 CVPDRQQYQGRLLAVTTHGLPCLAWASAQAQAKALSKHQDENSAVQLVENFCRNPDDDEEGV 272
QY 25 CFWDNGHLYREDQTSPAPGLRCLNWLDQSQS-GLASSAP-VSGAGN--HSYCRNPDEDPRG 80
Db 273 WCYYAGKPG 281
QY 81 WCYVSGEAG 89
RESULT 15
ENTRY A47136 #type complete
TITLE macrophage-stimulating protein 1 precursor - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 03-May-1994 #sequence_revision 14-Nov-1997 #text_change
11-Sep-1998
ACCESSION A40331; B40331; A47136; A61395
REFERENCE A40331
#authors Han, S.; Stuart, L.A.; Degen, S.J.F.
#journal Biochemistry (1991) 30:9768-9780
#title Characterization of the DNF15S2 locus on human chromosome 3:
identification of a gene coding for four kringle domains
with homology to hepatocyte growth factor.
#cross-references MUID:920002016
#accession A40331
##molecule_type DNA
#cross-references GB:M74179
#accession B40331
##molecule_type mRNA
##residues 1-711 ##label HAL
#cross-references GB:M74179
#accession B40331
##molecule_type mRNA
##residues 1-711 ##label HA2
#cross-references GB:M74178; NID:g183976; PID:g183977
#cross-references GB:M74178; NID:g183976; PID:g183977
REFERENCE A47136

```

Search completed: Fri Sep 17 18:26:39 1999
 Job time : 50 secs.

```

#authors Yoshimura, T.; Yuhki, N.; Wang, M.H.; Skeel, A.; Leonard,
E.J.
#journal J. Biol. Chem. (1993) 268:15461-15468
#title Cloning, sequencing, and expression of human macrophage
stimulating protein (MSP, MST1) confirms MSP as a member of
the family of kingle proteins and locates the MSP gene on
chromosome 3.
#cross-references MUID:93340141
#accession A47136
#molecule_type mRNA
##residues 1-12,'C','14-622,'F',624-711 ##label YOS
##Cross-references GB:L11924; NID:g398037; PID:g398038
##note authors translated the codon TTT for residue 623 as Leu;
parts of this sequence were determined by protein
sequencing

REFERENCE A61395
#authors Skeel, A.; Yoshimura, T.; Showalter, S.D.; Tanaka, S. ;
Appella, E.; Leonard, E.J.
#journal J. Exp. Med. (1991) 173:1227-1234
#title Macrophage stimulating protein: purification, partial amino
acid sequence, and cellular activity.
#cross-references MUID:91217635
#accession A61395
#molecule_type protein
##residues 230-247;288-291,'E',293-295,'X',297-301,'X',303,'E',305,
'EX',308-310;326-331,'X',333-341;484-501;530-536,'X',
538-546,'X',548-549,'SL';574-592,'E',594-596;602-611
##label SKE
##experimental_source plasma

GENETICS
#gene GDB:MST1; D3E15S2; DNF15S2; HGFL
. ##cross-references GDB:128833; OMIM:142408
#map_position 3p21-3p21.3
COMPLEX disulfide-bonded heterodimer of chains derived from the same
precursor

CLASSIFICATION #superfamily hepatocyte growth factor; kringle homology;
trypsin homology
KEYWORDS duplication; glycoprotein; growth factor; kringle; plasma
FEATURE
 1-18 #domain signal sequence #status predicted #label SIG\
 19-483 #product macrophage-stimulating protein 1 #status
predicted #label MAT \
 110-186 #domain alpha chain #status predicted #label ACH \
 191-268 #domain kringle homology #label KR1 \
 283-361 #domain kringle homology #label KR2 \
 370-448 #domain kringle homology #label KR3 \
 484-711 #domain beta chain #status predicted #label KR4 \
 56-78 #domain trypsin homology #label TRY \
 110-186,131-169,
157-181,191-268,
212-251,240-263,
283-361,304-343,
332-355,370-448,
391-431,419-443,
507-523,602-667,
632-646,657-685
#disulfide_bonds #status predicted \
72,296,615 #binding_site carbohydrate (Asn) (covalent) #status
predicted
SUMMARY #length 711 #molecular-weight 80379 #checksum 5521

Query Match 8.8%; Score 166; DB 1; Length 711;
Best Local Similarity 32.5%; Pred. No. 7.58e-11;
Matches 27; Conservative 17; Mismatches 35; Indels 4; Gaps 4;
Db 110 CIMNNGVGYRGTMATTVGGLPCQAWSHKFPNDHKYPTLRLNGLEENFCRNPDGDGGPWC 169
| : : : | :: | : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 25 CFWDNGHLYREDQTSPAPGLRCLNWLDA-QSGLASASAPVSGAG-NHSYCRNPDEDPRGPWC 82
Db 170 YTT-DPAVRF-QSCCGIKSCREAA 190
| : : : | :: | : : | : | : :
Qy 83 YVSGEAGVPEKRPCEDLRCPETT 105

```

```
*****
|   \   / |   \   / |   \   / |   \   / |
|     \ / |     \ / |     \ / |     \ / |
|       \ |       \ |       \ |       \ |
|         \|         \|         \|         \| (TM)
|           \ |           \ |           \ |
|             \|             \|             \| *****
*****
```

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPSEARCH_PP protein - protein database search, using Smith-Waterman algorithm
 Run on: Fri Sep 17 18:26:56 1999; MasPar time 8.89 seconds
 835.926 Million cell updates/sec
 Tabular output not generated.

Title: >US-09-084-491A-2
 Description: (1-263) from US09084491A.pep
 Perfect Score: 1883
 Sequence: 1 MLLAWVQAFLVSNMLAEAY PVDPQEGSTPLMGOAGTPGA 263

Scoring table: PAM 150

Gap 11

Searched: 77977 seqs, 28268293 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries
 Database: swiss-prot37
 1:swissprot

Statistics: Mean 46.647; Variance 77.329; scale 0.603

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description	Pred. No.
1	204	10.8	169	1	PLMN_RAT	PLASMINOGEN (EC 3.4.21)	1.10e-20
2	184	9.8	812	1	PLMN_MOUSE	PLASMINOGEN PRECURSOR	1.20e-16
3	183	9.7	4548	1	APOA_HUMAN	APOLIPOPROTEIN(A) PREC	1.89e-16
4	181	9.6	655	1	HGFL_HUMAN	HEPATOCYTE GROWTH FACT	4.70e-16
5	177	9.4	716	1	HGFL_MOUSE	HEPATOCYTE GROWTH FACT	2.88e-15
6	176	9.3	810	1	PLMN_ERIEU	PLASMINOGEN PRECURSOR	4.53e-15
7	173	9.2	810	1	PLMN_MACMU	PLASMINOGEN PRECURSOR	1.75e-14
8	171	9.1	790	1	PLMN_PIG	PLASMINOGEN (EC 3.4.21)	4.27e-14
9	172	9.1	812	1	PLMN_BOVIN	PLASMINOGEN PRECURSOR	2.73e-14
10	168	8.9	728	1	HGF_MOUSE	HEPATOCYTE GROWTH FACT	1.62e-13
11	167	8.9	728	1	HGF_HUMAN	HEPATOCYTE GROWTH FACT	2.53e-13
12	165	8.8	622	1	TRB_HUMAN	PROTHROMBIN PRECURSOR	6.12e-13
13	166	8.8	711	1	HGFL_HUMAN	HEPATOCYTE GROWTH FACT	3.94e-13
14	165	8.8	728	1	HGF_RAT	HEPATOCYTE GROWTH FACT	6.12e-13
15	164	8.7	625	1	TRB_BOVIN	PROTHROMBIN PRECURSOR	9.51e-13
16	163	8.7	1420	1	APOA_MACMU	APOLIPOPROTEIN(A) (EC	1.48e-12
17	161	8.6	810	1	PLMN_HUMAN	PLASMINOGEN PRECURSOR	3.54e-12
18	158	8.4	603	1	FA12_CAVPO	COAGULATION FACTOR XII	1.31e-11
19	156	8.3	431	1	URTB_DESRO	SALIVARY PLASMINOGEN A	3.11e-11
20	156	8.3	477	1	URT2_DESRO	SALIVARY PLASMINOGEN A	3.11e-11
21	155	8.2	566	1	URT_BOVIN	TISSUE PLASMINOGEN ACT	4.79e-11
22	152	8.1	333	1	PLMN_CANFA	PLASMINOGEN (EC 3.4.21)	1.73e-10
23	152	8.1	433	1	UROK_PAPCY	UROKINASE-TYPE PLASMIN	1.73e-10

ALIGNMENTS

PRT; 1.78e-08

STANDARD;

PRT; 1.78e-08

COAGULATION FACTOR XII

PLASMINOGEN (EC 3.4.21)

SALIVARY PLASMINOGEN A

UROKINASE-TYPE PLASMIN

PROTHROMBIN PRECURSOR

UROKINASE-TYPE PLASMIN

TISSUE PLASMINOGEN ACT

FAT

PLASMINOGEN (EC 3.4.21)

CDS-DIACYLGLYCEROL--SE

GENERAL SECRETION PATH

PLG.

RATTUS NORVEGICUS (RAT)

EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;

RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATITAE.

[1]

SEQUENCE FROM N.A.

RN

RP

RC

RX

RA

RT

RT

RL

J. BIOL. CHEM. 266:10825-10829 (1991).

-!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING

EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,

AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE

GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN

ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH

AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPDIN,

LAMININ AND VON WILLEBRAND FACTOR.

-!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN

ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO

FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.

-!- PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN IMMEDIATELY AFTER

DISSOCIATION FROM THE CLOT.

-!- SIMILARITY: CONTAINS 5 KRINGLE REGIONS.

-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY. BELONGS TO THE PLASMINOGEN SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch).

EMBL; M62832; G554488; -

DR PIR; A40522; A40522.

DR PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.

DR PROSITE; PS00135; TRYPSIN_SER; PARTIAL.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS50070; KRINGLE_2; 2.
 DR PFAM; PF00051; kringle; 1.
 DR HSSP; P00747; 1PMK.
 KW HYDROLASE; SERINE PROTEASE; PLASMA; GLYCOPROTEIN; FIBRINOLYSIS;
 KW TISSUE REMODELING; BLOOD COAGULATION; KRINGLE.
 FT NON_TER 1 1
 FT DOMAIN <1 10 KRINGLE 3 (BY SIMILARITY).
 FT DOMAIN 34 112 KRINGLE 4 (BY SIMILARITY).
 FT DOMAIN 139 >169 KRINGLE 5 (BY SIMILARITY).
 FT DISULFID 34 112 BY SIMILARITY.
 FT DISULFID 55 95 BY SIMILARITY.
 FT DISULFID 83 107 BY SIMILARITY.
 FT NON_TER 169 169 BY SIMILARITY.
 SQ SEQUENCE 169 AA; 18401 MW; B55004D7 CRC32;

Query Match 10.8%; Score 204; DB 1; Length 169;
 Best Local Similarity 26.2%; Pred. No. 1.10e-20;
 Matches 33; Conservative 44; Mismatches 40; Indels 9; Gaps 8;

Db 34 CYQNGKSYRGTSSTNTGKKCQSWSMTPHSKTPANFPDSGLEMMYCRNPDNDQRGP 93
 |: ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Qy 25 CFWDNGHLYREDQTSPPAPGLRCLNW--LDAQS-GLASAPVSGAG-NHSYCRNPDDEDPRGP 80

Db 94 WCFTT-DPSVR-WEYCNLKRCSSETGG-GVAE-SAIYPQVPSAPGTSETDCMYNGNGKEYRG 149
 ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Qy 81 WCYVSGEAGVPERKRPCTEDLRCPETTSQALPAFTTEIQEASEGPGADEVQ-VFAPANALPA 139

Db 150 KTAVTA 155
 :: :| 140 RSEAAA 145

Qy 140 RSEAAA 145

RESULT 2
 ID PLMN_MOUSE STANDARD; PRT; 812 AA.
 AC P20918;
 DT -01-FEB-1991 (REL. 17, CREATED)
 DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
 DE PLASMINOGEN PRECURSOR (EC 3.4.21.7).
 GN PLG.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
 OC RODENTIA; SCIURGNATHI; MURIDAE; MURINAE; MUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 91184812.
 RA DEGEN S.J., BELL S.M., SCHAEFER L.A., ELLIOTT R.W.;
 "Characterization of the cDNA coding for mouse plasminogen and
 localization of the gene to mouse chromosome 17.";
 RL GENOMICS 8:49-61(1990).

-!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
 A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
 EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
 AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
 GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
 ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
 AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
 LAMININ AND VON WILLEBRAND FACTOR.

-!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
 ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
 FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
 -!- PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN IMMEDIATELY AFTER
 DISSOCIATION FROM THE CLOT.

-!- IN THE PRESENCE OF THE INHIBITOR, THE ACTIVATION INVOLVES ONLY
 CLEAVAGE AFTER ARG-581, RESULTING IN 2 CHAINS HELD TOGETHER BY 2
 DISULFIDE BONDS. WITHOUT THE INHIBITOR, THE ACTIVATION INVOLVES
 ALSO REMOVAL OF THE ACTIVATION PEPTIDE.

-!- SIMILARITY: CONTAINS 5 KRINGLE REGIONS.
 -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 TRYPSIN FAMILY. BELONGS TO THE PLASMINOGEN SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; J04766; G200403; -.
 DR PIR; A38514; A38514.
 DR MGD; MGI:97620; PLG.
 DR PROSITE; PS00021; KRINGLE_1; 4.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR PROSITE; PS50070; KRINGLE_2; 5.
 DR PFAM; PF00051; kringle; 5.
 DR PFAM; PF00089; trypsin; 1.
 DR HSSP; P00747; 1PMK.
 KW HYDROLASE; SERINE PROTEASE; PLASMA; GLYCOPROTEIN; FIBRINOLYSIS;
 KW TISSUE REMODELING; BLOOD COAGULATION; KRINGLE; ZYMOGEN; SIGNAL.
 FT SIGNAL 1 19
 FT CHAIN 20 812 PLASMINOGEN.
 FT CHAIN 20 581 HEAVY CHAIN A.
 FT PEPTIDE 20 96 ACTIVATION PEPTIDE.
 FT CHAIN 97 581 SHORT FORM OF CHAIN A.
 FT CHAIN 582 812 LIGHT CHAIN B.
 FT DOMAIN 103 181 KRINGLE 1.
 FT DOMAIN 184 262 KRINGLE 2.
 FT DOMAIN 275 352 KRINGLE 3.
 FT DOMAIN 377 454 KRINGLE 4.
 FT DOMAIN 481 560 KRINGLE 5.
 FT DOMAIN 582 812 CATALYTIC.
 FT ACT_SITE 624 624 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 667 667 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 762 762 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 49 73 BY SIMILARITY.
 FT DISULFID 53 61 BY SIMILARITY.
 FT DISULFID 103 181 BY SIMILARITY.
 FT DISULFID 124 164 BY SIMILARITY.
 FT DISULFID 152 176 BY SIMILARITY.
 FT DISULFID 185 262 BY SIMILARITY.
 FT DISULFID 188 316 BY SIMILARITY.
 FT DISULFID 206 245 BY SIMILARITY.
 FT DISULFID 234 257 BY SIMILARITY.
 FT DISULFID 275 352 BY SIMILARITY.
 FT DISULFID 296 335 BY SIMILARITY.
 FT DISULFID 324 347 BY SIMILARITY.
 FT DISULFID 377 454 BY SIMILARITY.
 FT DISULFID 398 437 BY SIMILARITY.
 FT DISULFID 426 449 BY SIMILARITY.
 FT DISULFID 481 560 BY SIMILARITY.
 FT DISULFID 502 543 BY SIMILARITY.
 FT DISULFID 531 555 BY SIMILARITY.
 FT DISULFID 568 687 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 578 586 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 609 625 BY SIMILARITY.
 FT DISULFID 701 768 BY SIMILARITY.
 FT DISULFID 731 747 BY SIMILARITY.
 SQ SEQUENCE 812 AA; 90846 MW; 6C120F7A CRC32;
 Query Match 9.88; Score 184; DB 1; Length 812;
 Best Local Similarity 27.8%; Pred. No. 1.20e-16;
 Matches 35; Conservative 38; Mismatches 43; Indels 10; Gaps 9;

Db 377 CYQDGQSYRGTSSTTITGKKCQSWAAMPFRHSKTPENFPDAGLEMMYCRNPDGD-KGP 435
 |: ||| :: ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Qy 25 CFWDNGHLYREDQTSPPAPGLRCLNWLD--AQG-GLASAPVSGAG-NHSYCRNPDDEDPRGP 80

Db 436 WCYTT-DPSVR-WEYCNLKRCSSETGG-SVVELPTVSQEPS-GPSDSETDCMYNGKDYRG 491
 ||| :||| :: ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Qy 81 WCYVSGEAGVPEKRPCEDLRCPETTSQALPAFTTEIQEASEGPGADEVQ-VFAPPNALPA 139

RESULT	4
ID	HGFA_HUMAN STANDARD; PRT; 655 AA.
AC	Q04756; Q14726;
DT	01-JUN-1994 (REL. 29, CREATED)
DT	01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE	HEPATOCYTE GROWTH FACTOR ACTIVATOR PRECURSOR (EC 3.4.21.-) (HGF ACTIVATOR).
OS	HOMO SAPIENS (HUMAN).
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN	[1] SEQUENCE FROM N.A., AND PARTIAL SEQUENCE. TISSUE=LIVER, AND SERUM; MEDLINE; 93252878.
RA	MIYAZAWA K., SHIMOURA T., KITAMURA A., KONDO J., MORIMOTO Y., KITAMURA N.;
RA	"Molecular cloning and sequence analysis of the cDNA for a human serine protease responsible for activation of hepatocyte growth factor. Structural similarity of the protease precursor to blood coagulation factor XIII.";
RL	J. BIOL. CHEM. 268:10024-10028(1993).
RN	[2] SEQUENCE OF 40-655 FROM N.A.
RA	ZHAO S., ODELL C.;
RL	SUBMITTED (FEB-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC	-!- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY CONVERTING IT FROM A SINGLE CHAIN TO A HETERO-DIMERIC FORM.
CC	-!- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A DISULFIDE BOND.
CC	-!- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN PRECURSOR AND IS THEN ACTIVATED TO A HETERO-DIMERIC FORM.
CC	-!- TISSUE SPECIFICITY: LIVER.
CC	-!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC	-!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE-I DOMAIN.
CC	-!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE-II DOMAIN.
CC	-!- SIMILARITY: CONTAINS 1 KRINGLE REGION.
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS TRYPSIN FAMILY.
CC	-!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL ou the European Bioinformatics Institute. There are no restriction on use by non-profit institutions as long as its content is not modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch) or send an email to license@isb-sib.ch.
DR	EMBL; D14012; G219681; -.
DR	EMBL; Z69923; E225802; -.
DR	PIR; A46688; A46688.
DR	PROSITE; PS00021; KRINGLE_1; 1.
DR	PROSITE; PS00022; EGF_1; 2.
DR	PROSITE; PS00023; FIBRONECTIN_2; 1.
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.
DR	PROSITE; PS00135; TRYPSIN_SER; 1.
DR	PROSITE; PS01186; EGF_2; 1.
DR	PROSITE; PS01253; FIBRONECTIN_1; 1.
DR	PROSITE; PS50070; KRINGLE_2; 1.
PFAM	PF00008; EGF; 2.
PFAM	PF00039; fn1; 1.
PFAM	PF00040; fn2; 1.
PFAM	PF00051; kringle; 1.
PFAM	PF00089; trypsin; 1.
HSSP	P00763; 1DPO.
KW	HYDROLASE; GLYCOPROTEIN; PLASMA; SERINE PROTEASE; KRINGLE; SIGNAL
KW	EGF-LIKE DOMAIN; REPEAT; ZYMOGEN.
SIGNAL	1 30
PROPEP	31 372 CLEAVED IN ACTIVE FORM.
CHAIN	373 407 HEPATOCYTE GROWTH FACTOR ACTIVATOR CHAIN.

FT	CHAIN	408	655	HEPATOCYTE GROWTH FACTOR ACTIVATOR LONG CHAIN.
FT	DOMAIN	108	148	FIBRONECTIN TYPE- II.
FT	DOMAIN	160	198	EGF-LIKE 1.
FT	DOMAIN	200	240	FIBRONECTIN TYPE- I.
FT	DOMAIN	241	279	EGF-LIKE 2.
FT	DOMAIN	286	367	KRINGLE.
FT	DISULFID	408	655	CATALYTIC.
FT	ACT_SITE	447	447	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE	497	497	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE	598	598	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	DISULFID	108	133	BY SIMILARITY.
FT	DISULFID	122	148	BY SIMILARITY.
FT	DISULFID	164	175	BY SIMILARITY.
FT	DISULFID	169	186	BY SIMILARITY.
FT	DISULFID	188	197	BY SIMILARITY.
FT	DISULFID	202	230	BY SIMILARITY.
FT	DISULFID	228	237	BY SIMILARITY.
FT	DISULFID	245	256	BY SIMILARITY.
FT	DISULFID	250	267	BY SIMILARITY.
FT	DISULFID	269	278	BY SIMILARITY.
FT	DISULFID	286	367	BY SIMILARITY.
FT	DISULFID	307	349	BY SIMILARITY.
FT	DISULFID	338	362	BY SIMILARITY.
FT	DISULFID	394	521	INTRACHAIN (BY SIMILARITY).
FT	DISULFID	432	448	BY SIMILARITY.
FT	DISULFID	440	510	BY SIMILARITY.
FT	DISULFID	535	604	BY SIMILARITY.
FT	DISULFID	567	583	BY SIMILARITY.
FT	DISULFID	594	622	BY SIMILARITY.
FT	CARBOHYD	48	48	POTENTIAL.
FT	CARBOHYD	290	290	POTENTIAL.
FT	CARBOHYD	468	468	POTENTIAL.
FT	CARBOHYD	492	492	POTENTIAL.
FT	CARBOHYD	546	546	POTENTIAL.
FT	CONFLICT	644	644	R' -> Q (IN REF. 2).
SQ	SEQUENCE	655 AA;	70681 MW;	BFE0842D CRC32;
Query	Match	9.68;	Score 181;	DB 1; Length 655;
Best	Local Similarity	40.58;	Pred. No. 4.70e-16;	
Matches	30;	Conservative	19;	Mismatches 17; Indels 8; Gaps
Db	286	CFLGNGTGYRGVASTSASGLSCLAWNSDLLYQELHVDSVGAAALLGLGPHAYCRNPDNDE	345	
QY	25	CFWDNGHLYREDQTSPAPGLRCLNW---L--DAQ-SGLASAPVSGAGNHSYCRNPDDEDP	77	
Db	346	R-PWCYVVFKDSALS 358		
QY	78	RGPWCYVSGEAGVP 91		
RESULT	5			
ID	HGF_L_MOUSE	STANDARD;	PRT;	716 AA.
AC	P26928;			
DT	01-AUG-1992 (REL. 23, CREATED)			
DT	01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)			
DT	15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)			
DE	HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN PRECURSOR (MACROPHAGE STIMULATORY PROTEIN) (MSP).			
GN	HGFL			
OS	MUS MUSCULUS (MOUSE).			
RC	SEQUENCE FROM N.A. STRAIN=BALB/C; TISSUE=LIVER; MEDLINE; 92002017.			
RX	FRIEZNER DEGEN S.J., STUART L.A., HAN S., JAMISON C.S.;			
RA	"Characterization of the mouse cDNA and gene coding for a hepatocyte growth factor-like protein: expression during development."			
RT	BIOCHEMISTRY 30:9781-9791(1991).			
RL	- FUNCTION: PROBABLY HAS NO PROTEOLYTIC ACTIVITY, SINCE CRUCIAL AA			
CC	CHARACTERISTIC OF SERINE PROTEASES CATALYTIC SITES ARE NOT			

CC -!- CONSERVED.
 CC -!- TISSUE SPECIFICITY: LIVER. LOWER LEVELS IN LUNG, PLACENTA AND
 ADRENAL.
 CC -!- DEVELOPMENTAL STAGE: IS EXPRESSED AT LOW LEVELS DURING GESTATION.
 JUST BEFORE BIRTH THE LEVEL INCREASES DRAMATICALLY AND REMAINS
 STABLE AFTERWARDS.
 CC -!- PTM: MAY BE CLEAVED AFTER AA 488, TO YIELD A TWO-CHAIN MOLECULE
 HELD TOGETHER BY DISULFIDE BONDS, OR CLEAVED INTO TWO SEPARATE
 POLYPEPTIDES.

CC -!- SIMILARITY: CONTAINS 4 KRINGLE REGIONS.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 TRYPSIN FAMILY. BELONGS TO THE PLASMINOGEN SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; M74180; G193832; -.
 DR EMBL; M74181; G193834; -.
 DR MGD; MGI:96080; HGFL.
 DR PROSITE; PS000021; KRINGLE_1; 4.
 DR PROSITE; PS500070; KRINGLE_2; 4.
 DR PFAM; PF000051; kringle; 4.
 DR PFAM; PF000089; trypsin; 1.
 DR HSSP; P00747; 1PMK.

KW KRINGLE; GLYCOPROTEIN; SERINE PROTEASE HOMOLOG; SIGNAL.
 FT SIGNAL 1 31
 FT CHAIN 32 716
 FT DOMAIN 32 109
 FT DISULFID 56 78
 FT DISULFID 60 66
 FT DISULFID 60 186
 FT DISULFID 110 169
 FT DISULFID 131 181
 FT DISULFID 131 268
 FT DOMAIN 379 457
 FT DOMAIN 489 716
 FT DOMAIN 489 716
 FT DISULFID 56 78
 FT DISULFID 56 78
 FT DISULFID 110 186
 FT DISULFID 131 169
 FT DISULFID 157 181
 FT DISULFID 191 268
 FT DISULFID 191 333
 FT DISULFID 212 251
 FT DISULFID 240 263
 FT DISULFID 292 370
 FT DISULFID 313 352
 FT DISULFID 341 364
 FT DISULFID 379 457
 FT DISULFID 400 440
 FT DISULFID 428 452
 FT DISULFID 194 333
 FT DISULFID 212 251
 FT DISULFID 240 263
 FT DISULFID 292 370
 FT DISULFID 313 352
 FT DISULFID 341 364
 FT DISULFID 379 457
 FT DISULFID 400 440
 FT DISULFID 428 452
 FT DISULFID 194 333
 FT DISULFID 212 251
 FT DISULFID 512 528
 FT DISULFID 607 672
 FT DISULFID 637 651
 FT DISULFID 662 690
 FT CARBOHYD 72 72
 FT CARBOHYD 173 173
 FT CARBOHYD 305 305
 FT CARBOHYD 620 620
 FT CONFLICT 19 19
 SQ SEQUENCE 716 AA; 80588 MW; EDDAD87F CRC32;

Query Match 9.4%; Score 177; DB 1; Length 716;
 Best Local Similarity 41.3%; Pred. No. 2.88e-15;
 Matches 26; Conservative 9; Mismatches 26; Indels 2; Gaps 2;

DB 110 CIMDNGVSYRGTVARTAGGLPCQAWSRRFPNDHKYTPKNGLEENCRNPDGDPRGFWC 169
 | : ||| ||| : | | | : | : :||||| |||||
 QY 25 CFWDNGHLYREDQTSAPPGLRCLNWLD-QSGLASAPVSSGAG-NHSYCRNPDDEPRGPFWC 82

Db 170 YTT 172
 QY 83 YVS 85

RESULT 6
 ID PLMN_ERIEU STANDARD; PRT; 810 AA.
 ID PLMN_ERIEU STANDARD; PRT; 810 AA.
 AC Q29485;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE PLASMINOGEN PRECURSOR (EC 3.4.21.7).
 GN PLG.
 OS ERINACEUS EUROPAEUS (WESTERN EUROPEAN HEDGEHOG).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTERIA;
 OC INSECTIVORA; ERINACEIDAE; ERINACEINAE; ERINACEUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-LIVER;
 RX MEDLINE; 96025778.
 RA LAWN R.M.; BOONMARK N.W., SCHWARTZ K., LINDAHL G.E., WADE D.P.,
 BYRNE C.D., FONG K.J., MEER K., PATTHY L.;
 RT "The recurring evolution of lipoprotein(a). Insights from cloning of the hedgehog apolipoprotein(a)." J. BIOL. CHEM. 270:24004-24009(1995).
 RN [2]
 RP REVISIONS.
 RA LAWN R.M.;
 RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
 CC -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION, AND INFILTRATION; IN OVULATION IT WEAKENS THE WALLS OF THE GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN, LAMININ AND VON WILLEBRAND FACTOR.
 CC -!- ENZYME REGULATION: CONVERTED INTO PLASMINOGEN ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
 CC -!- PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
 CC -!- SIMILARITY: CONTAINS 5 KRINGLE REGIONS.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY. BELONGS TO THE PLASMINOGEN SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; U33171; G1806583; -.
 DR PROSITE; PS00021; KRINGLE_1; 5.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR PROSITE; PS50070; KRINGLE_2; 5.
 DR PFAM; PF00051; kringle; 5.
 DR PFAM; PF00089; trypsin; 1.
 DR HSSP; P00747; 1PMK.
 KW HYDROLASE; SERINE PROTEASE; PLASMA; GLYCOPROTEIN; FIBRINOLYSIS; TISSUE REMODELING; BLOOD COAGULATION; KRINGLE; ZYMOGEN; SIGNAL; SIGNAL 1 19
 FT CHAIN 20 810 PLASMINOGEN.
 FT CHAIN 20 582 HEAVY CHAIN A (BY SIMILARITY).
 FT CHAIN 583 810 LIGHT CHAIN B (BY SIMILARITY).
 FT DOMAIN 583 810 CATALYTIC.
 FT DOMAIN 103 181 KRINGLE 1.
 FT DOMAIN 185 262 KRINGLE 2.
 FT DOMAIN 275 352 KRINGLE 3.
 FT DOMAIN 379 456 KRINGLE 4.

PLMN_PIG	STANDARD;	PRT;	790 AA.
P06867;			
01-JAN-1988	(REL. 06, CREATED)		
01-FEB-1991	(REL. 17, LAST SEQUENCE UPDATE)		
01-OCT-1996	(REL. 34, LAST ANNOTATION UPDATE)		
PLASMINOGEN	(EC 3.4.21.7).		
SUS SCROFA (PIG).			
EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;			
ARTIODACTyla; SUIFORMES; SUINA; SUIDAE; SUS.			
[1] SEQUENCE OF 1-560.			
SCHALLER J., MARTI T., ROESSELET S.J., KAEMPFER U., RICKLI E.E.;			
"Amino acid sequence of the heavy chain of porcine plasmin. Comparison of the carbohydrate attachment sites with the human and bovine species."			
FIBRINOLYSIS 1:91-102(1987).			
[2] SEQUENCE OF 450-790.			
MEDLINE; 85203907.			
MARTI T., SCHALLER J., RICKLI E.E.;			
"Determination of the complete amino-acid sequence of porcine miniplasminogen."			
EUR. J. BIOCHEM. 149:279-285(1985).			
-!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION, AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN, LAMININ AND VON WILLEBRAND FACTOR.			
-!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.			
-!- PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.			
-!- SIMILARITY: CONTAINS 5 KRINGLE REGIONS.			
-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY. BELONGS TO THE PLASMINOGEN SUBFAMILY.			
PIR; A25834; A25834.			
PIR; S03733; S03733.			
PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.			
PROSITE; PS00021; KRINGLE_1; 5.			
PROSITE; PS00135; TRYPSIN_SER; 1.			
PROSITE; PS50070; KRINGLE_2; 5.			
PFAM; PF00051; kringle; 5.			
PFAM; PF00089; trypsin; 1.			
HSSP; P00747; 5HPG.			
KW HYDROLASE; SERINE PROTEASE; PLASMA; GLYCOPROTEIN; FIBRINOLYSIS; TISSUE REMODELING; BLOOD COAGULATION; KRINGLE; ZYMOGEN.			
KW TISSUE REMODELING; BLOOD COAGULATION; KRINGLE; ZYMOGEN.			
FT CHAIN 1 560 . HEAVY CHAIN A.			
FT CHAIN 561 790 . LIGHT CHAIN B.			
FT DOMAIN 561 790 . CATALYTIC.			
FT DOMAIN 84 162 . KRINGLE 1.			
FT DOMAIN 166 243 . KRINGLE 2.			
FT DOMAIN 256 333 . KRINGLE 3.			
FT DOMAIN 358 435 . KRINGLE 4.			
FT DOMAIN 461 540 . KRINGLE 5.			
FT ACT-SITE 602 602 . CHARGE RELAY SYSTEM.			
FT ACT-SITE 645 645 . CHARGE RELAY SYSTEM.			
FT ACT-SITE 740 740 . CHARGE RELAY SYSTEM.			
FT CARBOHYD 249 249 . CHARGE RELAY SYSTEM.			
SEQUENCE 790 AA; 88592 MW; EE597814 CRC32;			
Query Match 9 . Score 171; DB 1; Length 790;			
Best Local Similarity 26.4%; Pred. No. 4.27e-14;			
Matches 34; Conservative 36; Mismatches 49; Indels 10; Gaps 9;			
Db 358 CYRNGESYRTTSSTTGRKCQSWVSMTPHRHEKTPGNFPNAGLTMYCRNPDA-KSP 416			
: :: : : : : : : : : : : : : : : : :			
25 CFWDNGHLYREDQTSPAPGLRCLNW--LDAQS-GLASAPVSGAG-NHSYCRNPDEPRGP 80			
QY 417 WCYTT-DPRVR-WEYCNKKCSET-EQQVTINFPAIAQVPSVEDLSED-CMFNGNKRYRGK 472			
DR PROSITE; PS00021; KRINGLE_1; 5.			
DR EMBL; X79402; G494963; -.			
DR PROSITE; PS00134; TRYPSIN_HIS; 1.			
DR EMBL; K02935; G163552; -.			
DR PROSITE; PS00135; TRYPSIN_SER; 1.			
DR PROSITE; PS50070; KRINGLE_2; 5.			
DR PIR; A25835; PLBO.			
DR PROSITE; PS00051; kringle; 5.			
CC or send an email to license@isb-sib.ch).			

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR PFAM; PF00089; trypsin; 1.

DR HSSP; P00747; 2PK4.

KW HYDROLASE; SERINE PROTEASE; PLASMA; GLYCOPROTEIN; FIBRINOLYSIS; TISSUE REMODELING; BLOOD COAGULATION; KRINGLE; ZYMOGEN; SIGNAL.

FT SIGNAL 1 26

FT CHAIN 27 812 PLASMINOGEN.

FT CHAIN 27 583 HEAVY CHAIN A.

FT DOMAIN 584 812 LIGHT CHAIN B.

FT DOMAIN 584 812 CATALYTIC.

FT DOMAIN 110 188 KRINGLE 1.

FT DOMAIN 192 269 KRINGLE 2.

FT DOMAIN 282 359 KRINGLE 3.

FT DOMAIN 384 461 KRINGLE 4.

FT DOMAIN 485 564 KRINGLE 5.

FT CARBOHYD 315 315 GROWTH FACTOR; KRINGLE; GLYCOPROTEIN; SERINE PROTEASE HOMOLOG;

FT CARBOHYD 365 365 SIGNAL; ALTERNATIVE SPLICING.

FT ACT_SITE 624 624 CHARGE RELAY SYSTEM.

FT ACT_SITE 667 667 CHARGE RELAY SYSTEM.

FT ACT_SITE 762 762 CHARGE RELAY SYSTEM.

FT CONFLICT 335 335 N->D (IN REF. 2).

FT CONFLICT 516 516 Q->H (IN REF. 2).

FT CONFLICT 555 555 P->L (IN REF. 2).

FT CONFLICT 744 744 T->R (IN REF. 3).

SQ SEQUENCE 812 AA; 91216 MW; 3E3C0328 CRC32;

Query Match 9.1%; Score 172; DB 1; Length 812;

Best Local Similarity 31.1%; Pred. No. 2.73e-14;

Matches 33; Conservative 25; Mismatches 39; Indels 9; Gaps 8;

Db 384 CYHGNQSYRGTSSTTITGRKCQSWSMTPPHRLKTPENYPNAGLTMNYCRNPAD-KSP 442

QY 25 CFWDNGHLYREDQTSPAPGLRCLNW-LDAQSGL-ASAPVSGAG-NHSYCRNPDEDPRGP 80

Db 443 WCYTT-DPRVR-WEFCNLKKCSETPEQ-VPA-APQAPGVENPPEAD 484

QY 81 WCYVSGEAGVPEKRPCEDLRCPETTSQALPAFFTEIQEASEGPGAD 126

Query Match 8.9%; Score 168; DB 1; Length 728;

Best Local Similarity 31.9%; Pred. No. 1.62e-13;

Matches 30; Conservative 23; Mismatches 34; Indels 7; Gaps 5;

Db 387 SSGQDCYRGNGKNYMGNLSKTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNKNYCRNPDD 446

QY 21 GSG-GCFWDNGHLYREDQTSPAPGLRCLNW--LDA-QSGLASAPVSGAGNHSYCRNPDE 75

Db 447 DAHGPCWCY-TGNPLIP-WDYCPISRCGEGTPTI 478

QY 76 DPRGPWCYVSGEAGVPEKRPCETTSQALPAFFTEIQEASEGPGAD 109

RESULT 10

ID HGF_MOUSE STANDARD; PRT; 728 AA.

AC Q08048;

DT 01-NOV-1995 (REL. 32, CREATED)

DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)

DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)

DE HEPATOCYTE GROWTH FACTOR PRECURSOR (SCATTER FACTOR) (SF)

DE (HEPATOPOEITIN A).

GN HGF.

OS MUS MUSCULUS (MOUSE).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;

OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.

[1] SEQUENCE FROM N.A., AND SEQUENCE OF 496-504.

TRN TISSUE=MAMMARY FIBROBLAST;

RP MEDLINE; 94183257.

RC RA SASAKI M., NISHIO M., SASAKI T., ENAMI J.;

"Identification of mouse mammary fibroblast-derived mammary growth factor as hepatocyte growth factor";

BIOCHEM. BIOPHYS. RES. COMMUN. 199:772-779(1994).

-!- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL HEPATOCYTE CELLS, SEEMS TO BE AN HEPATOTROPHIC FACTOR, AND ACTS AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.

-!- IT HAS NO DETECTABLE PROTEASE ACTIVITY.

CC -SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A DISULFIDE BOND.

CC -!- ALTERNATIVE PRODUCTS: A SHORT FORM OF HGF IS PRODUCED BY ALTERNATIVE RNA SPLICING. THE SEQUENCE SHOWN HERE IS THAT OF THE LONG FORM.

CC -!- SIMILARITY: CONTAINS 4 KRINGLE REGIONS.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY. BELONGS TO THE PLASMINOGEN SUBFAMILY.

CC DR EMBL; D10212; G220436; -.

DR EMBL; D10213; G220438; -.

DR MGI; 96079; HGF.

DR PROSITE; PS00021; KRINGLE_1; 4.

DR PROSITE; PS50070; KRINGLE_2; 4.

DR PFAM; PF00051; kringle; 4.

DR PFAM; PF00089; trypsin; 1.

DR HSPP; P14210; 2HGF.

KW GROWTH FACTOR; KRINGLE; GLYCOPROTEIN; SERINE PROTEASE HOMOLOG;

KW SIGNAL; ALTERNATIVE SPLICING.

FT SIGNAL 1 32 BY SIMILARITY.

FT CHAIN 33 495 HEPATOCYTE GROWTH FACTOR ALPHA CHAIN.

FT CHAIN 496 728 HEPATOCYTE GROWTH FACTOR BETA CHAIN.

FT MOD_RES 33 33 PYRROLIDONE CARBOXYLIC ACID (BY SIMILARITY).

FT DOMAIN 33 128 PAP.

FT DOMAIN 129 207 KRINGLE 1.

FT DOMAIN 212 289 KRINGLE 2.

FT DOMAIN 306 384 KRINGLE 3.

FT DOMAIN 392 470 KRINGLE 4.

FT DOMAIN 496 728 SERINE PROTEASE-LIKE.

FT DISULFID 71 97 BY SIMILARITY.

FT DISULFID 75 85 BY SIMILARITY.

FT DISULFID 488 607 INTERCHAIN (BY SIMILARITY).

FT CARBOHYD 295 295 POTENTIAL.

FT CARBOHYD 403 403 POTENTIAL.

FT CARBOHYD 569 569 POTENTIAL.

FT CARBOHYD 656 656 POTENTIAL.

FT VARSPLIC 163 167 MISSING (IN SHORT FORM).

SQ SEQUENCE 728 AA; 82944 MW; 967CDF1E CRC32;

Query Match 8.9%; Score 168; DB 1; Length 728;

Best Local Similarity 31.9%; Pred. No. 1.62e-13;

Matches 30; Conservative 23; Mismatches 34; Indels 7; Gaps 5;

Db 387 SSGQDCYRGNGKNYMGNLSKTRSGLTCSMWDKNMEDLHRHIFWEPDASKLNKNYCRNPDD 446

QY 21 GSG-GCFWDNGHLYREDQTSPAPGLRCLNW--LDA-QSGLASAPVSGAGNHSYCRNPDE 75

Db 447 DAHGPCWCY-TGNPLIP-WDYCPISRCGEGTPTI 478

QY 76 DPRGPWCYVSGEAGVPEKRPCETTSQALPAFFTEIQEASEGPGAD 109

RESULT 11

ID HGF_HUMAN STANDARD; PRT; 728 AA.

AC P14210;

DT 01-JAN-1990 (REL. 13, CREATED)

DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)

DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)

DE HEPATOCYTE GROWTH FACTOR PRECURSOR (SCATTER FACTOR) (SF)

DE (HEPATOPOEITIN A).

GN HGF OR HPTA.

OS HOMO SAPIENS (HUMAN).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;

OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.

[1] RN SEQUENCE FROM N.A.

RP SEQUENCE FROM N.A.

RC MEDLINE; 91340155.

RA SEKI T., HAGIYA M., SHIMONISHI M., NAKAMURA T., SHIMIZU S.;

RT "Organization of the human hepatocyte growth factor-encoding gene.";

RL GENE 102:213-219(1991).

[2] RN SEQUENCE FROM N.A.

RP TISSUE=PLACENTA;

RC MEDLINE; 89392017.

RA MIYAZAWA K., TSUBOUCHI H., NAKA D., TAKAHASHI K., OKIGAKI M..

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation.

RA ARAKAKI N., NAKAYAMA H., HIRONO S., SAKIYAMA O., TAKAHASHI K.,
 RA GOHDA E., DAIKUHARA Y., KITAMURA N.;
 RT "Molecular cloning and sequence analysis of cDNA for human hepatocyte
 growth factor.";
 RT BIOCHEM. BIOPHYS. RES. COMMUN. 163:967-973(1989).
 RL [3]

RN SEQUENCE FROM N.A.
 RP TISSUE=LEUKOCYTE;
 RC MEDLINE; 91025062.
 RX SEKI T., IHARA I., SUGIMURA A., SHIMONISHI M., NISHIZAWA T.,
 RA ASAMI O., HAGIYA M., NAKAMURA T., SHIMIZU S.;
 RT "Isolation and expression of cDNA for different forms of hepatocyte
 growth factor from human leukocyte.";
 RT BIOCHEM. BIOPHYS. RES. COMMUN. 172:321-327(1990).
 RL [4]

RN SEQUENCE FROM N.A., AND SEQUENCE OF 55-73 AND 495-520.
 RP TISSUE=LIVER;
 RC MEDLINE; 90066676.
 RX NAKAMURA T., NISHIZAWA T., HAGIYA M., SEKI T., SHIMONISHI M.,
 RA SUGIMURA A., TASHIRO K., SHIMIZU S.;
 RT "Molecular cloning and expression of human hepatocyte growth factor.";
 RL NATURE 342:440-443(1990).
 RN [5]

RN SEQUENCE FROM N.A.
 RP TISSUE=EMBRYONIC FIBROBLAST;
 RC MEDLINE; 911334393.
 RX WEIDNER K.M., ARAKAKI N., HARTMANN G., VANDEKERCKHOVE J., WEINGART S.,
 RA RIEDER H., FONATSCH C., TSUBOUCHI H., HISHIDA T., DATIKUHARA Y.,
 RA BIRCHMEIER W.;
 RT "Evidence for the identity of human scatter factor and human
 hepatocyte growth factor.";
 RT native and recombinant human hepatocyte growth factor.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 88:7001-7005(1991).
 RN [6]

RN SIGNAL SEQUENCE CLEAVAGE SITE.
 RP MEDLINE; 91207365.
 RA YOSHIIYAMA Y., ARAKAKAI N., NAKA D., TAKAHASHI K., HIRONO S., KONDO J.,
 RA NAKAYAMA H., GOHDA E., KITAMURA N., TSUBOUCHI H., ISHII T.,
 RA HISHIDA T., DAIKUHARA Y.;
 RT "Identification of the N-terminal residue of the heavy chain of both
 native and recombinant human hepatocyte growth factor.";
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 175:660-667(1991).
 RN [7]

RN CARBOHYDRATE-BINDING SITE 476.
 RP MEDLINE; 93129192.
 RA SHIMIZU N., HARA H., SOGABE T., SAKAI H., IHARA I., INOUE H.,
 RA NAKAMURA T., SHIMIZU S.;
 RT "Hepatocyte growth factor is linked by O-glycosylated oligosaccharide
 on the alpha chain.";
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 189:1329-1335(1992).
 RN [8]

RN MUTAGENESIS.
 RP MEDLINE; 92331602.
 RA LOKKER N.A., MARK M.R., LUIS E.A., BENNETT G.L., ROBBINS K.A.,
 RA BAKER J.B., GODOWSKI P.J.;
 RT "Structure-function analysis of hepatocyte growth factor:
 RT identification of variants that lack mitogenic activity yet retain
 RT high affinity receptor binding.";
 RL EMBO J.11:2503-2510(1992).
 RN [9]

RN STRUCTURE BY NMR OF 31-127.
 RP MEDLINE; 98154323.
 RA ZHOU H., MAZZULLA M.J., KAUFMAN J.D., STAHL S.J., WINGFIELD P.T.,
 RA RUBIN J.S., BOTIARO D.P., BYRD R.A.;
 RT "The solution structure of the N-terminal domain of hepatocyte growth
 RT factor reveals a potential heparin-binding site.";
 RL STRUCTURE 6:109-116(1998).
 RN [10]

RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 35-210.
 RP MEDLINE; 99036858.
 RA ULTSCH M., LOKKER N.A., GODOWSKI P.J., DE VOS A.M.;
 RT "Crystal structure of the NK1 fragment of human hepatocyte growth
 factor at 2.0-A resolution";
 RL STRUCTURE 6:1383-1393(1998).
 RL

CC -!- FUNCTION: HGF IS A POTENT MITOGEN FOR MATURE PARENCHYMAL
 CC HEPATOCYTE CELLS, SEEKS TO BE AN HEPATOTROPHIC FACTOR, AND ACTS
 CC AS GROWTH FACTOR FOR A BROAD SPECTRUM OF TISSUES AND CELL TYPES.
 CC IT HAS NO DETECTABLE PROTEASE ACTIVITY.
 CC -!- SUBUNIT: DIMER OF AN ALPHA CHAIN AND A BETA CHAIN LINKED BY A
 CC DISULFIDE BOND.
 CC -!- SIMILARITY: CONTAINS 4 KRINGLE REGIONS.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY. BELONGS TO THE PLASMINOGEN SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC DR EMBL; D90334; G219700;
 CC DR EMBL; D90318; G219700; JOINED.
 CC DR EMBL; D90319; G219700; JOINED.
 CC DR EMBL; D90320; G219700; JOINED.
 CC DR EMBL; D90322; G219700; JOINED.
 CC DR EMBL; D90323; G219700; JOINED.
 CC DR EMBL; D90324; G219700; JOINED.
 CC DR EMBL; D90325; G219700; JOINED.
 CC DR EMBL; D90326; G219700; JOINED.
 CC DR EMBL; D90327; G219700; JOINED.
 CC DR EMBL; D90328; G219700; JOINED.
 CC DR EMBL; D90329; G219700; JOINED.
 CC DR EMBL; D90330; G219700; JOINED.
 CC DR EMBL; D90331; G219700; JOINED.
 CC DR EMBL; D90332; G219700; JOINED.
 CC DR EMBL; D90333; G219700; JOINED.
 CC DR EMBL; M29145; G306846;
 CC DR EMBL; M60718; G184032;
 CC DR EMBL; X16323; G32082;
 CC DR EMBL; M73239; G337936;
 CC DR EMBL; M73240; G337938;
 CC PIR; JH0579; JH0579;
 CC PIR; S06794; S06794;
 CC PDB; 2HGF; 24-JUN-98.
 CC PDB; 1BHT; 18-NOV-98.
 DR MIM; 142409;
 DR PROSITE; PS00021; KRINGLE_1; 4.
 DR PROSITE; PS50070; KRINGLE_2; 4.
 DR PFAM; PF00051; kringle; 4.
 DR PFAM; PF00089; trypsin; 1.
 KW GROWTH FACTOR; KRINGLE; GLYCOPROTEIN; SERINE PROTEASE HOMOLOG;
 KW SIGNAL; 3D-STRUCTURE;
 FT SIGNAL 1 31
 FT CHAIN 32 494
 FT CHAIN 495 728
 FT MOD_RES 32 32
 FT DOMAIN 32 127
 FT DOMAIN 128 206
 FT DOMAIN 211 288
 FT DOMAIN 305 383
 FT DOMAIN 391 469
 FT DOMAIN 495 728
 FT DISULFID 70 96
 FT DISULFID 74 84
 FT DISULFID 128 206
 FT DISULFID 149 189
 FT DISULFID 177 201
 FT DISULFID 487 604
 FT CARBOHYD 294 294
 FT CARBOHYD 402 402
 FT CARBOHYD 476 476
 FT CARBOHYD 566 566
 FT CARBOHYD 653 653
 FT CONFLICT 32 33
 FT CONFLICT 78 78
 K -> N (IN REF. 4).
 K -> N (IN REF. 4).

MEDLINE; 92378975.
 MIYATA T., ARUGA R., UMEYAMA H., BEZEAUD A., GUILLIN M.-C., IWANAGA S.; "Prothrombin Salakta: substitution of glutamic acid-466 by alanine reduces the fibrinogen clotting activity and the esterase activity."; BIOCHEMISTRY 31:7457-7462(1992). [17]

VARIANT TOKUSHIMA.
 MEDLINE; 87185407.
 MIYATA T., MORITA T., INOMOTO T., KAWAUCHI S., SHIRAKAMI A., IWANAGA S.; "Prothrombin Tokushima, a replacement of arginine-418 by tryptophan that impairs the fibrinogen clotting activity of derived thrombin Tokushima"; BIOCHEMISTRY 26:1117-1122(1987). [18]

VARIANT TOKUSHIMA.
 MEDLINE; 87101511.
 INOMOTO T., SHIRAKAMI A., KAWAUCHI S., SHIGEKIYO T., SAITO S., MIYOSHI K., MORITA T., IWANAGA S.; "Prothrombin Tokushima: characterization of dysfunctional thrombin derived from a variant of human prothrombin."; BLOOD 69:565-569(1987). [19]

VARIANT TOKUSHIMA.
 MEDLINE; 92256895.
 IWAHANA H., YOSHIMOTO K., SHIGEKIYO T., SHIRAKAMI A., SAITO S., ITAKURA M.; "Detection of a single base substitution of the gene for prothrombin Tokushima. The application of PCR-SSCP for the genetic and molecular analysis of dysprothrombinemia."; INT. J. HEMATOL. 55:93-100(1992). [20]

VARIANT TYPE-3.
 MEDLINE; 83204687.
 BOARD P.G., SHAW D.C.; "Determination of the amino acid substitution in human prothrombin type 3 (157 Glu leads to Lys) and the localization of a third thrombin cleavage site."; BR. J. HAEMATOL. 54:245-254(1983).

-!- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII, AND, IN COMPLEX WITH THROMBOMODULIN, PROTEIN C.
 -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
 -!- TISSUE SPECIFICITY: SYNTHESIZED IN THE LIVER; FOUND IN PLASMA.
 -!- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS, RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSOMAL ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUE ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSIO OF PROTHROMBIN TO THROMBIN.
 -!- DISEASE: DEFECTS IN F2 ARE THE CAUSE OF VARIOUS FORMS OF DYSPROTHROMBINEMIA.

-!- PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN & FACTORS VA & XA IN CA-DEPENDENT INTERACTIONS; FACTOR XA REMOVES THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT & HEAVY CHAINS. THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF THROMBIN.
 -!- IT IS NOT KNOWN WHETHER 1 OR 2 SMALLER ACTIVATION PEPTIDES, WITH ADDITIONAL CLEAVAGE AFTER 314-ARG, ARE RELEASED IN NATURAL BLOOD CLOTTING.

-!- THROMBIN CAN ITSELF CLEAVE THE AMINO TERMINAL FRAGMENT (FRAGMENT 1) OF THE PROTHROMBIN, PRIOR TO ITS ACTIVATION BY FACTOR XA.
 -!- THE CLEAVAGE AFTER R-198, OBSERVED IN VITRO, DOES NOT OCCUR IN PLASMA.

-!- SIMILARITY: CONTAINS 2 KRINGLE REGIONS.
 -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M17262; G339641; - .
DR EMBL; V00595; E5121; - .
DR PIR; A00914; TBHU.
...: remainder of annotations omitted.

	Query Match	Score 8.88;	Score 165;	DB 1;	Length 622;
Best Local Matches	Similarity 37.78;	Pred. No. 6.12e-13;			
Matches 26;	Conservative 13;	Mismatches 26;	Indels 4;	Gaps 3;	
Db 213	CVPDRCQOYQGRLAVTTHGLPCLAWASAQAKALSKHODENSAVQLVENFCRNPDGDEEGV	272			
QY 25	CFWDNGHLYREDQTSPAPGLRCLNWLDQSQS-GLASAP-VSGAGN-HSYCRNPDDEDPRGP	80			
Db 273	WCYVAGKPG 281				
QY 81	: :				
QY 81	WCYVSGEAG 89				

RESULT 13
ID HGFL_HUMAN STANDARD; PRT; 711 AA.
AC P26927;
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN PRECURSOR (MACROPHAGE STIMULATORY PROTEIN) (MSP) (MACROPHAGE STIMULATING PROTEIN).
GN MST1 OR HGFL.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE; 92002016.
RA HAN S., STUART L.A., FRIEZNER DEGEN S.J.;
RT "Characterization of the DNF15S2 locus on human chromosome 3: identification of a gene coding for four kringle domains with homology to hepatocyte growth factor.";
RT BIOCHEMISTRY 30:9768-9780(1991).
RL 484
CC -!- FUNCTION: PROBABLY HAS NO PROTEOLYTIC ACTIVITY, SINCE CRUCIAL AA CONSERVED.
CC -!- PTM: MAY BE CLEAVED AFTER AA 484, TO YIELD A TWO-CHAIN MOLECULE HELD TOGETHER BY DISULFIDE BONDS, OR TWO SEPARATE POLYPEPTIDES.
CC -!- SIMILARITY: CONTAINS 4 KRINGLE REGIONS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY. BELONGS TO THE PLASMINOGEN SUBFAMILY.
CC
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
DR EMBL; M74178; G183977; - .
DR EMBL; M74179; G183979; - .
DR PIR; A40331; A40331.
DR MIM; 142408; - .
DR PROSITE; PS00021; KRINGLE_1; 4 .
DR PROSITE; PS50070; KRINGLE_2; 4 .
DR PFAM; PF00051; Kringle; 4 .
DR PFAM; PF00089; trypsin; 1 .
DR HSSP; P00763; 1SLW.
KW KRINGLE; GLYCOPROTEIN; SERINE PROTEASE HOMOLOG; SIGNAL; POLYMORPHISM.
KW

- F2. OS "Characterization of the bovine prothrombin gene.";
 EUKARYOTA; METAZOA; CHORDATA; MAMMALIA; EUTHERIA;
 ARTIODACTyla; RUMINANTIA; PECORA; BOVIDEA; BOVINAE; BOS.
 [1] SEQUENCE FROM N.A.
 RP MEDLINE; 88245190.
- RA IRWIN D.M., ROBERTSON K.A., MACGILLIVRAY R.T.A.;
 RT "Structure and evolution of the bovine prothrombin gene.";
 RL J. MOL. BIOL. 200:31-45(1988).
- [2] RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 84203525.
- RA MCGILLIVRAY R.T.A., DAVIE E.W.;
 RT "Characterization of bovine prothrombin mRNA and its translation product.";
 RL BIOCHEMISTRY 23:1626-1634(1984).
- [3] RN
 RP SEQUENCE OF 44-625, DISULFIDE BONDS, AND CARBOHYDRATE-BINDING SITES.
 RA MAGNUSSON S., SOTTRUP-JENSEN L., PETERSEN T.E., CLAEYS H.;
 RL (IN) HEMKER H.C., VELTKAMP J.J. (EDS.);
 BOERHAAVE SYMPOSIUM ON PROTHROMBIN AND RELATED COAGULATION FACTORS,
 PP.25-46, LEIDEN UNIVERSITY PRESS, LEIDEN (1975).
- [4] RN
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF ACTIVATION PEPTIDE 1.
 RX MEDLINE; 86296631.
 RA PARK C.H., TULINSKY A.;
 RT "Three-dimensional structure of the kringle sequence: structure of prothrombin fragment 1.";
 RT "Structure of bovine prothrombin fragment 1 refined at 2.25-A resolution.";
 RL BIOCHEMISTRY 25:3977-3982(1986).
- [5] RN
 RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF ACTIVATION PEPTIDE 1.
 RX MEDLINE; 91311686.
 RA SESHDARI T.-P., TULINSKY A., SKRZYPczAK-JANKUN E., PARK C.H.;
 RT "The Ca²⁺ ion and membrane binding structure of the Gla domain of Ca-prothrombin fragment 1";
 RL BIOCHEMISTRY 31:2554-2566(1992).
- [6] RN
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF ACTIVATION PEPTIDE 1.
 RX MEDLINE; 92190185.
 RA SORIANO-GARCIA M., PADMANABHAN K., DE VOS A.M., TULINSKY A.;
 RT "The Ca²⁺ ion and membrane binding structure of the Gla domain of human fibrinogen bound to bovine thrombin at 2.3-A resolution.";
 RL J. BIOL. CHEM. 267:7911-7920(1992).
- [7] RN
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RX MEDLINE; 92218459.
 RA BRANDSTETTER H., TURK D., HOEFFKEN H.W., GROSSE D., STUERZEBECHER J.,
 MARTIN P.D., ROBERTSON W., HUBER R., BODE W., EDWARDS B.F.P.;
 RT "The structure of residues 7-16 of the A alpha-chain of human fibrinogen bound to bovine thrombin at 2.3-A resolution.";
 RL J. BIOL. CHEM. 267:7911-7920(1992).
- [8] RN
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RX MEDLINE; 92389319.
 RA HOEFFKEN W., HUBER R.;
 RT "Refined 2.3 A X-ray crystal structure of bovine thrombin complexes formed with the benzamidine and arginine-based thrombin inhibitors NAPAP, 4-TAPAP and MQPA. A starting point for improving antithrombotics.";
 RT "The ornithodorin-thrombin crystal structure, a key to the TAP enigma?";
 RL EMBO J. 15:6011-6017(1996).
- [9] RN
 RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF COMPLEX WITH ORNITHODORIN.
 RA VAN DE LOCHT A., STUBBS M.T., BODE W., FRIEDRICH T., BOLLSCHWEILER C.,
 RA HOFFKEN W., HUBER R.;
 RT "The ornithodorin-thrombin crystal structure, a key to the TAP enigma?";
 RL EMBO J. 15:6011-6017(1996).
- [10] RN
 RP GENE STRUCTURE.
 RX MEDLINE; 86077733.
 RA IRWIN D.M., AHERN K.G., PEARSON G.D., MCGILLIVRAY R.T.A.;
- RT "Characterization of the bovine prothrombin gene.";
 RL BIOCHEMISTRY 24:6854-6861(1985).
 CC -!- FUNCTION: THROMBIN, WHICH CLEAVES BONDS AFTER ARG & LYS, CONVERTS FIBRINOGEN TO FIBRIN AND ACTIVATES FACTORS V, VII, VIII, XIII, XIV, AND, IN COMPLEX WITH THROMMODULIN, PROTEIN C.
 CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -!- TISSUE SPECIFICITY: SYNTHESIZED IN THE LIVER; FOUND IN PLASMA.
 CC -!- PTM: THE GAMMA-CARBOXYGLUTAMYL RESIDUES, WHICH BIND CALCIUM IONS, RESULT FROM THE CARBOXYLATION OF GLUTAMYL RESIDUES BY A MICROSONAL ENZYME, THE VITAMIN K-DEPENDENT CARBOXYLASE. THE MODIFIED RESIDUES ARE NECESSARY FOR THE CA-DEPENDENT INTERACTION WITH A NEGATIVELY CHARGED PHOSPHOLIPID SURFACE, WHICH IS ESSENTIAL FOR THE CONVERSION OF PROTHROMBIN TO THROMBIN.
 CC -!- PROTHROMBIN IS ACTIVATED ON THE SURFACE OF A PHOSPHOLIPID MEMBRANE THAT BINDS THE AMINO END OF PROTHROMBIN & FACTORS VA & XIa IN CA-DEPENDENT INTERACTIONS; FACTOR XIa REMOVES THE ACTIVATION PEPTIDE & CLEAVES THE REMAINING PART INTO LIGHT & HEAVY CHAINS.
 CC -!- THE ACTIVATION PROCESS STARTS SLOWLY BECAUSE FACTOR V ITSELF HAS TO BE ACTIVATED BY THE INITIAL, SMALL AMOUNTS OF THROMBIN.
 CC -!- THROMBIN CAN ITSELF CLEAVE THE AMINO TERMINAL FRAGMENT (FRAGMENT 1) OF THE PROTHROMBIN, PRIOR TO ITS ACTIVATION BY FACTOR XIa.
 CC -!- SIMILARITY: CONTAINS 2 KRINGLE REGIONS.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPSIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC DR EMBL; V00135; G808945;
 CC DR EMBL; J00041; G163755;
 CC DR PIR; A00915; TBB0.
 CC DR PIR; A00915; TBB0.
 CC DR PIR; S02537;
 CC DR PDB; 1BBR; 31-JAN-94.
 CC DR PDB; 1ETR; 31-JAN-94.
 CC DR PDB; 1ETS; 31-JAN-94.
 CC DR PDB; 1ETT; 31-JAN-94.
 CC DR PDB; 1HRT; 31-JAN-94.
 CC DR PDB; 2PF1; 31-JAN-94.
 CC DR PDB; 2PF2; 31-JAN-94.
 CC DR PDB; 2SPT; 31-MAY-94.
 CC DR PDB; 1MKW; 07-JUL-97.
 CC DR PDB; 1MKY; 07-JUL-97.
 CC DR PDB; 1TBO; 14-OCT-96.
 CC DR PDB; 1TBR; 14-OCT-96.
 CC DR PDB; 1TOC; 23-JUL-97.
 CC DR PDB; 1VIT; 21-APR-97.
 CC DR PDB; 1YCP; 06-MAY-98.
 CC DR PDB; 1AOH; 17-JUN-98.
 CC DR PROSITE; PS00011; GLU_CARBOXYLATION; 1.
 CC DR PROSITE; PS00021; KRINGLE_1; 2.
 CC DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC DR PROSITE; PS00135; TRYPSIN_SER; 1.
 CC DR PROSITE; PS50070; KRINGLE_2; 2.
 CC DR PFAM; PF00051; kringle; 2.
 CC DR PFAM; PF00089; trypsin; 1.
 CC DR PFAM; PF00594; gta; 1.
 CC FT PROPEP ? ?
 CC FT CHAIN 44 625 PROTHROMBIN.
 CC FT PEPTIDE 44 199 ACTIVATION PEPTIDE (FRAGMENT 1).
 CC FT PEPTIDE 200 317 ACTIVATION PEPTIDE (FRAGMENT 2).
 CC FT CHAIN 318 366 THROMBIN LIGHT CHAIN (A).
 CC FT CHAIN 367 625 THROMBIN HEAVY CHAIN (B).
 CC FT DOMAIN 109 187 KRINGLE 1.
 CC FT DOMAIN 214 292 KRINGLE 2.

SITE	199	CLEAVAGE (BY THROMBIN).
SITE	318	CLEAVAGE (BY FACTOR XA).
SITE	367	CLEAVAGE (BY FACTOR XA).
ACT_SITE	409	CHARGE RELAY SYSTEM.
ACT_SITE	465	CHARGE RELAY SYSTEM.
ACT_SITE	571	CHARGE RELAY SYSTEM.
MOD_RES	50	GAMMA-CARBOXYGLUTAMIC ACID.
MOD_RES	51	GAMMA-CARBOXYGLUTAMIC ACID.
MOD_RES	58	GAMMA-CARBOXYGLUTAMIC ACID.
MOD_RES	60	GAMMA-CARBOXYGLUTAMIC ACID.
MOD_RES	63	GAMMA-CARBOXYGLUTAMIC ACID.
MOD_RES	64	GAMMA-CARBOXYGLUTAMIC ACID.
MOD_RES	69	GAMMA-CARBOXYGLUTAMIC ACID.
MOD_RES	70	GAMMA-CARBOXYGLUTAMIC ACID.
MOD_RES	73	GAMMA-CARBOXYGLUTAMIC ACID.
MOD_RES	76	GAMMA-CARBOXYGLUTAMIC ACID.
CARBOHYD	120	120
CARBOHYD	144	144
CARBOHYD	419	419
DISULFID	61	66
DISULFID	91	104
DISULFID	109	187
DISULFID	130	170
DISULFID	158	182
DISULFID	214	292
DISULFID	235	275
DISULFID	263	287
DISULFID	339	485
DISULFID	394	410
DISULFID	539	553
DISULFID	567	597
VARIANT	600	600
CONFILCT	231	231
CONFILCT	249	249
CONFILCT	288	288
CONFILCT	353	353
CONFILCT	355	355
CONFILCT	549	550
HELIX	50	52
HELIX	57	62
HELIX	68	73
TURN	74	75
HELIX	79	89
TURN	90	93
HELIX	99	105
TURN	106	107
STRAND	110	110
END	112	112

Query Match Score 164; DB 1;
 Best Local Similarity Length 625;
 Pred. No. 9.51e-13;

b	274	WCYVADQPG	282
v	81	WCYVSGEAG	89

search completed: Fri Sep 17 18:27:21 1999
job time : 25 secs.

Best Available Copy